Journal of Scientific Computing (2019) 78:625–652
https://doi.org/10.1007/s10915-018-0783-x

V-cycle Multigrid Algorithms for Discontinuous Galerkin Methods on Non-nested Polytopic Meshes

P. F. Antonietti[1] · **G. Pennesi**[1]

Received: 12 September 2017 / Revised: 4 June 2018 / Accepted: 9 July 2018 / Published online: 20 July 2018

Abstract
In this paper we analyze the convergence properties of V-cycle multigrid algorithms for the numerical solution of the linear system of equations stemming from discontinuous Galerkin discretization of second-order elliptic partial differential equations on polytopic 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 algorithms 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 minimum number of smoothing steps, which depends on p, is chosen sufficiently large.

Keywords Discontinuous Galerkin · Polygonal grids · Multi-level methods · V-cycle · Non-nested spaces

Mathematics Subject Classification 65F10 · 65M55 · 65N22

1 Introduction

The discontinuous Galerkin (DG) method was introduced in 1973 by Reed and Hill for the discretization of the neutron equation [59]. Then, DG methods have been proposed to deal with elliptic and parabolic problems: some of the most relevant earlier works include Baker [15], Wheeler [65] and Arnold [12], whose contributions put the basis for the development of the interior penalty DG methods. In the last forty years the scientific and industrial community has shown a growing interest in DG methods—see for example [36,37,52,60] and the references therein for an overview. On the one hand, the features of DG methods have been

This work has been supported by the research grant PolyNuM founded by Fondazione Cariplo and Regione Lombardia, and by the SIR Project No. RBSI14VT0S funded by MIUR.

✉ G. Pennesi
giorgio.pennesi@polimi.it

P. F. Antonietti
paola.antonietti@polimi.it

[1] MOX-Laboratory for Modelling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy

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 mesh element, 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. [3–6,8,10,17–19,34,49,57,66]. In particular, the efficient approach presented in [34] is based on defining a local polynomial discrete space by making use of the bounding box of each element [48]: this technique together with a careful choice of the discontinuity penalization parameter allow for polytopic elements that can be characterized by faces of arbitrarily small measure and, as shown in [32], see also [8], possibly by an unbounded number of faces.

The development of fast solvers and preconditioners for the linear system of equations stemming from (high-order) DG discretizations has been an intensive research area in recent years. A recent strand of the literature has focused on Schwarz domain decomposition methods, cf. Ref. [1,2,6,9,31,41,42,44–46,54,64], and two-level and multigrid/multilevel techniques, cf. Ref. [8,11,14,18,19,28–30,38,50]. The efficiency of those methods can be further improved in the case of polygonal grids, because the flexibility of the element shape couples very well with the possibility of defining agglomerated meshes, which is the key ingredient for the development of multigrid algorithms. In [8] a two-level scheme and W-cycle multigrid methods are analyzed to solve the linear system of equations arising from high-order DG discretizations on polytopic grids. One iteration of the proposed methods consists of an iterative application of the smoothing Richardson operator and a recursive subspace correction step. In particular, the latter is based on a nested sequence of discontinuous discrete polynomial spaces, where the underlying polytopic grids are defined by agglomeration. While being perfectly suited for multilevel schemes, the process of element agglomeration might feature itself some limitations. Indeed, agglomeration leads to coarser grids with an increasing number of faces and this might affect the conditioning of the coarser components of the solver and the overall efficiency.

In this paper we aim at overcoming this issue by analyzing a multigrid 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 procedure. 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 and where the discrete bilinear forms are chosen differently on each level, cf. Ref. [20,69,70]. The first non-nested multilevel method was introduced by Bank and Dupont in [16]; a generalized framework was developed by Bramble, Pasciak and Xu in [26], and then widely used in the analysis of non-nested multigrid iterations, cf. Ref. [21–25,27,51,61,67,68]. The method of [26], 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 the bilinear forms 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 [43], which permits to develop a full analysis of V-cycle multigrid methods in a non-nested framework relax-

ing the hypothesis $\mathcal{A}_j(I_{j-1}^j u, I_{j-1}^j u) \leq \mathcal{A}_{j-1}(u,u)$. We prove that our V-cycle scheme with non-nested spaces converges uniformly with respect to the discretization parameters, namely the mesh size h and the polynomial degree p, provided that the number of smoothing steps, which depends on p, is chosen sufficiently large. This result extends the theory of [8] where W-cycle multigrid methods for high-order DG methods with nested spaces have been 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 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 our 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. Finally, in Sect. 7 we draw some conclusions.

2 Model Problem and Its DG Discretization

We consider the weak formulation of the Poisson problem, subject to homogeneous Dirichlet boundary conditions: find $u \in V = H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$\mathcal{A}(u,v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x \qquad \forall v \in V, \tag{1}$$

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 polytopic 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 $h_j = \max_{\kappa \in \mathcal{T}_j} h_\kappa$. For the sake of simplicity, we assume that on each level the mesh $\mathcal{T}_j$ is quasi-uniform. 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 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 $x \eqsim y$ 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 non-nested hp-multigrid schemes. This method is based on employing a set of non-nested polytopic 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 constraint

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

We also assume that the polynomial degree varies 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 neighboring elements. This implies that, for $d = 2$, a *face* always consists of a line segment, whereas for $d = 3$, the *faces* of $\mathcal{T}_j$ are general shaped polygons. Thereby, we assume that each face of an element $\kappa \in \mathcal{T}_j$ can be subdivided into a set of co-planar $(d-1)$-dimensional simplexes and we refer to them as *faces*. In order to introduce the DG formulation, it is helpful to distinguish between boundary and interior 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 $F \in \mathcal{F}_j^I$, $F \subset \partial\overline{\kappa}^{\pm}$, where $\kappa^{\pm}$ are two adjacent elements in $\mathcal{T}_j$. Furthermore, we denoted by $\mathcal{F}_j = \mathcal{F}_j^I \cup \mathcal{F}_j^B$ 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 simplexes is given. Moreover, for the forthcoming analysis, we require that the following assumptions hold, cf. [32,33].

Assumption 1 For any $j = 1, \ldots, J$, given $\kappa \in \mathcal{T}_j$ there exists a set of non-overlapping d-dimensional simplexes $T_l \subset \kappa$, $l = 1, \ldots, n_\kappa$, 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, $\cup_{l=1}^{n_\kappa} \overline{T_l} \subset \overline{\kappa}$, 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 simplexes $\{\mathfrak{s}_i\}_{i=1}^{m_\kappa}$, 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 For each $\mathcal{T}_j$, $j = 1, \ldots, J$, we assume that there exists a covering $\mathcal{T}_j^{\#} = \{\mathcal{S}_\kappa\}_\kappa$ of $\mathcal{T}_j$ consisting of shape-regular d-dimensional simplexes $\mathcal{S}_\kappa$, such that, for any $\kappa \in \mathcal{T}_j$, there is $\mathcal{S}_\kappa \in \mathcal{T}_j^{\#}$ satisfying $\kappa \subset \mathcal{S}_\kappa$ and $h_{\mathcal{S}_K} := diam(\mathcal{S}_K) \lesssim h_\kappa$. We also assume that for any $j = 1, \ldots, J$, it holds

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

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

Remark 3 Assumption 1 allows to employ very general polygonal and polyhedral elements. Indeed, if $\kappa \in \mathcal{T}_j$ is a polygonal/polyhedral element and $F \subset \partial\kappa$ is one of its faces, then Assumption 1 allows the size of F to be small compared to the diameter h_κ of κ, provided that the height of the related simplex T_l, with base F, is comparable to h_κ. We refer to [32] for more details.

2.2 DG Formulation

In order to introduce the DG discretization of (1), we first 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 is shared by $\kappa^{\pm} \in \mathcal{T}_j$, 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 $\boldsymbol{\tau}$ and v on F from $\kappa^{\pm}$, respectively. The jump and average operators across F are then defined as follows:

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

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

If $F \in \mathcal{F}_j^B$ is a boundary face, we set accordingly $\{\!\{\boldsymbol{\tau}\}\!\} = \boldsymbol{\tau}$, $[\![v]\!] = v\,\mathbf{n}$, cf. [13]. Let $\mathcal{R}_j : [L^1(\mathcal{F}_j)]^d \to [V_j]^d$ be the lifting operator defined as

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

cf. [13].
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

$$\begin{aligned}\mathcal{A}_j(u,v) = &\sum_{\kappa \in \mathcal{T}_j} \int_\kappa \Big[\nabla u \cdot \nabla v + \mathcal{R}_j([\![u]\!]) \cdot \nabla v + \mathcal{R}_j([\![v]\!]) \cdot \nabla u\Big] \mathrm{d}x \\ &+ \sum_{F \in \mathcal{F}_j} \int_F \sigma_j [\![u]\!] \cdot [\![v]\!]\, \mathrm{d}s, \end{aligned} \tag{5}$$

where, according to [39,40], $\sigma_j \in L^\infty(\mathcal{F}_j)$ denotes the interior penalty stabilization function, which is defined by

$$\sigma_j(x) = C_\sigma^j \frac{p_j^2}{\{h_\kappa\}_H}, \quad x \in F \in \mathcal{F}_j, \tag{6}$$

with $C_\sigma^j > 0$ independent of p_j, $|F|$ and $|\kappa|$, and where $\{\cdot\}_H$ is the harmonic average given by

$$\{h_\kappa\}_H = \begin{cases} \dfrac{2h_{\kappa^+} h_{\kappa^-}}{h_{\kappa^+} + h_{\kappa^-}}, & F \in \mathcal{F}_j^I,\ \overline{F} \subset \partial\overline{\kappa^+} \cap \partial\overline{\kappa^-}, \\ h_\kappa, & F \in \mathcal{F}_j^B,\ F \subset \partial\overline{\kappa} \cap \partial\overline{\Omega}. \end{cases}$$

Remark 4 Formulation (5) is based on the lifting operators $\mathcal{R}_j$ and allows to introduce the discrete gradient operator $\mathcal{G}_j : V_j \to [V_j]^d$, defined as

$$\mathcal{G}_j(v) = \nabla_j v + \mathcal{R}_j([\![v]\!]) \quad \forall\, j = 1, \ldots, J, \tag{7}$$

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

$$\mathcal{A}_J(u_J, v_J) = \int_\Omega f v_J\, \mathrm{d}x \quad \forall v_J \in V_J. \tag{8}$$

3 Preliminary Results

In this section we recall some preliminary results that 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. Let $\kappa \in \mathcal{T}_j$, the following bound holds*

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

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

The proof of Lemma 1 combines Assumption 1 with the idea of [37, Proof of Lemma 1.49].

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

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

We refer to [32] for the proof.

We endow each discrete space $V_j,\ j = 1, \ldots, J$, with the following DG norm:

$$\|w\|^2_{DG,j} = \sum_{\kappa\in\mathcal{T}_j}\int_\kappa |\nabla w|^2\,\mathrm{d}x + \sum_{F\in\mathcal{F}_j}\int_F \sigma_j|[\![w]\!]|^2\,\mathrm{d}s. \tag{9}$$

The well-posedness of the DG formulation (8) on each level $j = 1, \ldots, J$ is established in the following lemma.

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

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

The second bound holds provided that the constants $C^j_\sigma,\ j = 1, \ldots, J$, appearing in (6) are chosen sufficiently large.

Next, we recall the following approximation result, which is an analogous bound presented in [34, Theorem 5.2]. This result exploits the properties of the extension operator $\mathcal{E} : H^s(\Omega) \to H^s(\mathbb{R}^d),\ s \in \mathbb{N}_0$, such that $\mathcal{E}v|_\Omega = v$ and $\|\mathcal{E}v\|_{H^s(\mathbb{R}^d)} \lesssim \|v\|_{H^s(\Omega)}$, introduced in [62].

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

$$\sum_{\kappa\in\mathcal{T}_j}\|v - \tilde{\Pi}_j v\|_{H^q(\kappa)} \lesssim \sum_{\kappa\in\mathcal{T}_j}\frac{h_\kappa^{s-q}}{p_\kappa^{k-q}}\|\mathcal{E}v\|_{H^k(\mathcal{S}_\kappa)}, \quad \text{for} \quad 0 \leq q \leq k,$$

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

Remark 5 From Lemma 4, for uniform orders $p_\kappa = p_j$ and $h_\kappa = h_j \ \forall j = 1, \ldots, J$, we point out that, if $v \in H^k(\Omega)$, the following bound follows:

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

The result presented in Lemma 4 leads to the following error bounds for the underlying interior penalty DG scheme, which follows from the energy norm error bounds that have been proved in [34], see also [32] in the general case. The corresponding L^2-estimates can be found in [8].

Corollary 1 *Assume that Assumptions 1 and 4 hold. Let $u_j \in V_j$, $j = 1, \ldots, J$, be the DG solution of problem (8) posed on level j,* i.e.,

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

If the solution u of (1) is sufficiently regular, i.e. $u \in H^k(\Omega) \cap H_0^1(\Omega)$, $k \ge 2$, 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 $p_j \ge 1$.

Remark 6 We point out that the bounds in Corollary 1 are optimal in h and suboptimal 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 [47] for quadrilateral meshes providing the solution belongs to a suitable augmented Sobolev space. The issue of proving optimal estimates as the ones in [47] on polytopic meshes is an open problem and it is under investigation. In the following, we will write:

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

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

We also need to introduce an appropriate inverse inequality, cf. [8].

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*

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

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 [7] for a similar result on standard grids, and to [8] for its extension to polygonal grids.

Theorem 5 *Let Assumptions 1, 2 and 3 be satisfied. 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 of [26] for multigrid methods with 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 [11], 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.

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

$$(A_j w, v)_{L^2(\Omega)} = \mathcal{A}_j(w, v) \quad \forall w, v \in V_j, \quad j = 1, \ldots, J, \tag{10}$$

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

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

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. $I_{j-1}^j : V_{j-1} \to V_j$

$$(I_{j-1}^j v_H, w_h)_{L^2(\Omega)} = (v_H, w_h)_{L^2(\Omega)} \; \forall w_h \in V_j, \tag{11}$$

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 $P_j^{j-1} : V_j \to V_{j-1}$

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

According to (10), problem (8) can be written in the following equivalent form: find $u_J \in V_J$ such that

$$A_J u_J = f_J, \tag{12}$$

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 the 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 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

$$A_j z = g, \tag{13}$$

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 (13) obtained after one iteration of the non-nested V-cycle scheme with initial guess $z_0 \in V_j$ and m_1, m_2 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 exactly.

Algorithm 1 Multigrid V-cycle iteration for the solution of problem (12)

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

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 define G_j^* as its adjoint with respect to $\mathcal{A}_j(\cdot, \cdot)$. Following [43], we make three standard assumptions in order to prove convergence of Algorithm 1:

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

$$|\!|\!| (\mathrm{Id}_j - I_{j-1}^{j} P_j^{j-1}) v_h |\!|\!|_{1,j} \leq C_Q |\!|\!| v_h |\!|\!|_{1,j} \qquad \forall v_h \in V_j, \quad j = 2, \ldots, J.$$

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

$$\left| \mathcal{A}_j((\mathrm{Id}_j - I_{j-1}^{j} P_j^{j-1}) v_h, v_h) \right| \leq C_A \frac{|\!|\!| v_h |\!|\!|_{2,j}^2}{\Lambda_j} \qquad \forall v_h \in V_j, \quad j = 2, \ldots, J.$$

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

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

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

The convergence analysis of the V-cycle method is stated in the following theorem that gives an estimate for the error propagation operator related to the j-th level iteration with m_1 and m_2 pre- and post-smoothing steps, respectively. The error propagation operator is defined as

$$\begin{cases} \mathbb{E}_{1,m_1,m_2} v &= 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 1, \\ \mathbb{E}_{j,m_1,m_2} v &= \mathbb{G}^*_{j,m_2}(\mathrm{Id}_j - I^j_{j-1} P^{j-1}_j + I^j_{j-1}\mathbb{E}_{j-1,m_1,m_2} P^{j-1}_j)\mathbb{G}_{j,m_1} v, \quad j > 1, \end{cases}$$

where $\mathbb{G}_{j,m} = (G_j)^m$ and $\mathbb{G}^*_{j,m} = (G^*_j)^m$, $m \geq 1$.

Theorem 6 *If Assumptions* **A.1**, **A.2** *and* **A.3** *hold, then*

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

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

We refer to [43] for the proof of Theorem 6 in an abstract setting. In the following, we prove the validity of Assumptions **A.1**, **A.2** and **A.3** for our algorithm. We start with a two-level approach, i.e. $J = 2$, and we consider the two-level method for the solution of (8), 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 Validity 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_{J-1}. In the following, we also consider the L^2-projection operator on the space V_J defined as

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

Remark 7 From the definition of I^J_{J-1} given in (11), it holds $I^J_{J-1} = Q_J|_{V_{J-1}}$.

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 [9]. Let $\mathcal{G}_j$ be the discrete gradient operator (7) introduced in Remark 4, and consider the following problem: $\forall v_j \in V_j$, find $\mathcal{H}(v_j) \in H^1_0(\Omega)$ such that

$$\int_\Omega \nabla \mathcal{H}(v_j) \cdot \nabla w \, \mathrm{d}x = \int_\Omega \mathcal{G}_j(v_j) \cdot \nabla w \, \mathrm{d}x \quad \forall w \in H^1_0(\Omega). \tag{14}$$

It is shown in [9] that $\mathcal{H}(\cdot)$ possesses good approximation properties in terms of providing an H^1-conforming approximant of the discontinuous function v_j:

Theorem 7 *Let* Ω *be a bounded convex polygonal/polyhedral domain in* $\mathbb{R}^d$, $d = 2, 3$. *Given* $v_j \in V_j$, *we write* $\mathcal{H}(v_j) \in H^1_0(\Omega)$ *to be the approximation defined in* (14). *Then, the following approximation and stability results hold:*

$$\|v_j - \mathcal{H}(v_j)\|_{L^2(\Omega)} \lesssim \frac{h_j}{p_j} \|\sigma_j^{\frac{1}{2}} [\![v_j]\!]\|_{L^2(\mathcal{F}_h)}, \qquad |\mathcal{H}(v_j)|_{H^1(\Omega)} \lesssim \|v_j\|_{DG,j}. \tag{15}$$

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 $C_{stab} = C_{stab}(p_J)$, independent of the mesh size such that*

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

Here $C_{stab}(p_J) = \mathcal{O}(p_J)$.

Proof Let $v_H \in V_{J-1}$, by the definition of the DG norm (9), we need to estimate:

$$\|I_{J-1}^J v_H\|_{DG,J}^2 = \|\nabla_J (I_{J-1}^J v_H)\|_{L^2(\mathcal{T}_J)}^2 + \|\sigma_J^{\frac{1}{2}} |[\![I_{J-1}^J v_H]\!]|\|_{L^2(\mathcal{F}_J)}^2. \tag{16}$$

We next bound each of the two terms on the right hand side separately. For the first one we let $\mathcal{H}_H = \mathcal{H}(v_H)$ be defined as in (14). Then:

$$\begin{aligned}\|\nabla_J (I_{J-1}^J v_H)\|_{L^2(\mathcal{T}_J)}^2 \leq\ & \|\nabla_J (I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 \\ & + \|\nabla_J (\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 + |\mathcal{H}_H|_{H^1(\Omega)}^2,\end{aligned} \tag{17}$$

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 (15):

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

In order to estimate the first term on the right hand side in (17) 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:

$$\|\nabla_J (I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H))\|_{L^2(\mathcal{T}_J)}^2 \lesssim p_J^4 h_J^{-2} \|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2. \tag{18}$$

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

$$\|I_{J-1}^J v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2 \lesssim \|I_{J-1}^J v_H - \mathcal{H}_H\|_{L^2(\mathcal{T}_J)}^2 + \|\mathcal{H}_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2. \tag{19}$$

Using Lemma 4 and the Poincaré inequality (since $\mathcal{H}_H \in H_0^1(\Omega)$) we have

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

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

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

Using Remark 7, the continuity of Q_J with respect to the L^2-norm, Lemma 4 and (15) we have

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

Thanks to the previous estimates and inequality (19), we obtain

$$\|I_{J-1}^{J} v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\mathcal{T}_J)}^2 \lesssim \frac{h_J^2}{p_J^2}\|v_H\|_{DG,J-1}^2. \tag{20}$$

The above estimate, together with (18), (17) and the second bound of (15) lead to

$$\|\nabla_J(I_{J-1}^{J} v_H)\|_{L^2(\mathcal{T}_J)}^2 \lesssim p_J^2 \, \|v_H\|_{DG,J-1}^2. \tag{21}$$

Next we bound the second term on the right hand side in (16). 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

$$\begin{aligned} &\|\sigma_J^{\frac{1}{2}} [\![I_{J-1}^{J} v_H]\!]\|_{L^2(\mathcal{F}_J)}^2 \lesssim \\ &\frac{p_J^2}{h_J} \sum_{\kappa \in \mathcal{T}_J} \Big(\|I_{J-1}^{J} v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\partial\kappa)}^2 + \|\widetilde{\Pi}_J(\mathcal{H}_H) - \mathcal{H}_H\|_{L^2(\partial\kappa)}^2 \Big), \end{aligned} \tag{22}$$

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

$$\|I_{J-1}^{J} v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\partial\kappa)}^2 \lesssim \frac{p_J^2}{h_J}\|I_{J-1}^{J} v_H - \widetilde{\Pi}_J(\mathcal{H}_H)\|_{L^2(\kappa)}^2. \tag{23}$$

To bound the second term on the right hand side in (22), we first exploit the continuous trace inequality on polygons of Lemma 1 with $\epsilon = p_J$, obtaining

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

then, by summing over $\kappa \in \mathcal{T}_J$, using the approximation property of Lemma 4 and the Poincaré inequality, we obtain

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

From the previous inequality and the bound (23), (22) becomes:

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

where we also used inequality (20). This estimate together with (21) leads to

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

where $\mathsf{C}_{\mathsf{stab}}(p_J) = \mathcal{O}(p_J)$. □

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* $C_Q \lesssim p_J^2$.

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

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

Similarly, it holds

$$\mathcal{A}_{J-1}(P_J^{J-1} v_h, P_J^{J-1} v_h) \lesssim p_J^2\ \mathcal{A}_J(v_h, v_h) \quad \forall v_h \in V_J. \tag{24}$$

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

$$\mathcal{A}_J(I_{J-1}^J P_J^{J-1} v_h, I_{J-1}^J P_J^{J-1} v_h) \lesssim p_J^2\ \mathcal{A}_{J-1}(P_J^{J-1} v_h, P_J^{J-1} v_h). \tag{25}$$

By adding and subtracting v_h to both arguments of $\mathcal{A}_J(\cdot,\cdot)$ on the left hand side of (25), and using (24) we obtain

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

that concludes the proof. □

4.3 Validity 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".

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

$$\|(\mathrm{Id}_J - I_{J-1}^J P_J^{J-1})v_J\|_{L^2(\Omega)} \lesssim \frac{h_J^2}{p_J^{2-\mu}} |\!|\!| v_J |\!|\!|_{2,J} \quad \forall v_J \in V_J. \tag{26}$$

Thanks to Lemma 7, it is possible to show the following result.

Theorem 8 *The regularity-approximation property* **A.2** *holds with* $C_A \lesssim p_J^{2+\mu}$, $\mu = 0, 1$.

Proof Theorem 5 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 $v \in V_J$:

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

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

and the proof is complete. □

4.4 Validity of Assumption A.3

Proposition 2 *Assumption* **A.3** *holds with* $C_R = \mathcal{O}(1)$.

Proof We have:

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

and so

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

We now prove that $\Big(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\Big)$ 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

$$\Lambda_J(\overline{u}, \overline{u})_{L^2(\Omega)} < \mathcal{A}_J(\overline{u}, \overline{u}). \tag{27}$$

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 satisfies $\mathcal{A}_J(\phi_i^J, \phi_j^J) = \lambda_i^J \delta_{ij}$, 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 (27) 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 $\Big(\mathrm{Id}_J - \frac{1}{\Lambda_J}A_J\Big)$ is a positive definite operator. □

Remark 8 We observe that, as we need to satisfy the condition $m > 2C_A C_R$ of Theorem 6, we can guarantee the convergence of the method based on employing a number of smoothing steps such that $m \gtrsim p_J^{2+\mu}$, which is in agreement to what proved for W-cycle algorithms in [11] and [8] in the case of nested grids.

Remark 9 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 $C_R = \mathcal{O}(1)$ also on the arbitrary levels j, $j-1$, because each level j satisfies Assumption **A.3** with constant $C_R^j = \mathcal{O}(1)$. Assumptions **A.2** and **A.1** are satisfied with $C_A = \max_j\{C_A^j\}$ and $C_Q = \max_j\{C_Q^j\}$, respectively, where C_A^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 6 and to demonstrate the capability of our algorithm also in practical cases. We focus on a two dimensional problem posed on the unit square $\Omega = (0, 1)^2$. For the simulations, we consider the sets of polygonal grids shown in Figure 1. Each polygonal mesh is generated by using the software package `PolyMesher` [63]. In particular the finest grids (Level 4) of Fig. 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 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 corresponding finer one.

First of all, we verify the estimate of Lemma 6, by numerically evaluating $\mathsf{C}_{\mathsf{stab}}(p)$, 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 $\mathsf{C}_{\mathsf{stab}}(p)$ as a function of p. Figure 2 show that, as expected, $\mathsf{C}_{\mathsf{stab}}(p)$ 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 Fig. 1 and numerically evaluate the constant δ_j in Theorem 6 based on selecting the Richardson smoother with

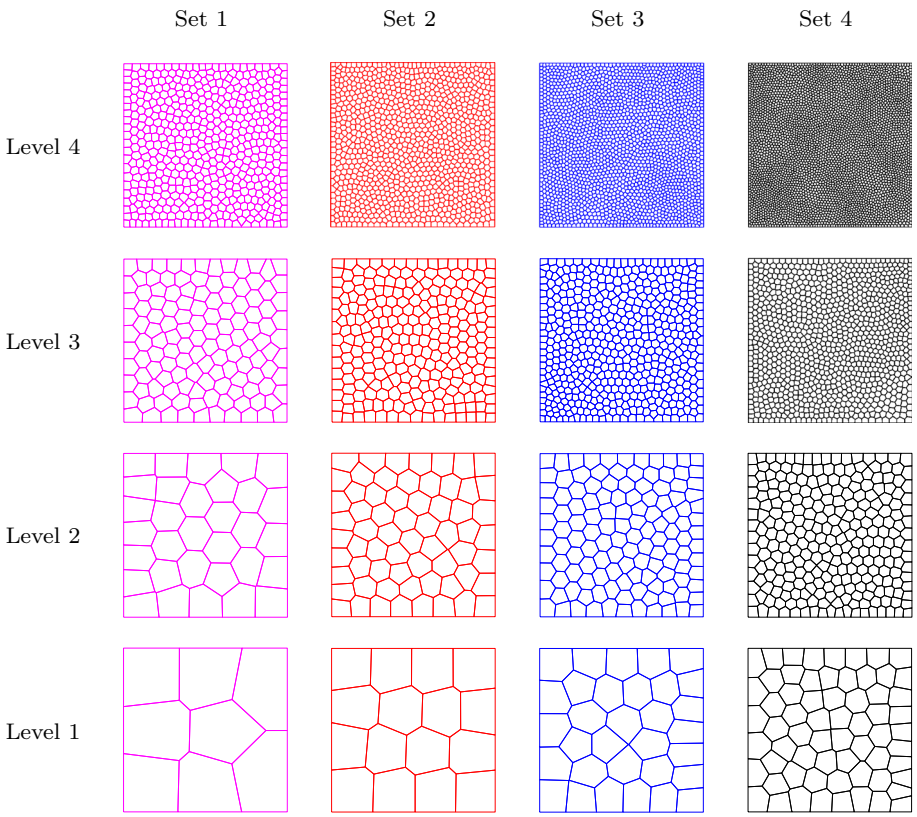

Fig. 1 Sets of non-nested grids employed for numerical simulations

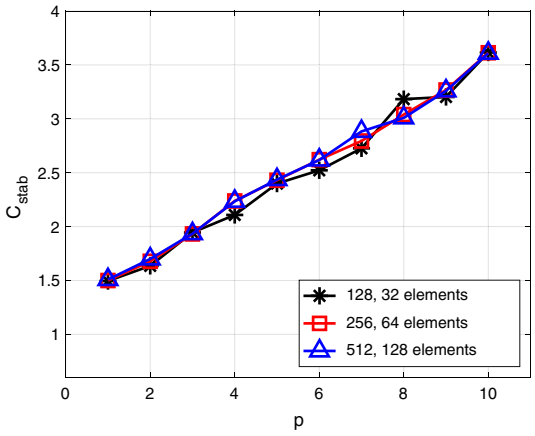

Fig. 2 Estimates of $\mathsf{C}_{\mathsf{stab}}(p)$ in Lemma 6 as a function of p for three couples of non-nested Voronoi meshes as shown in Fig. 1

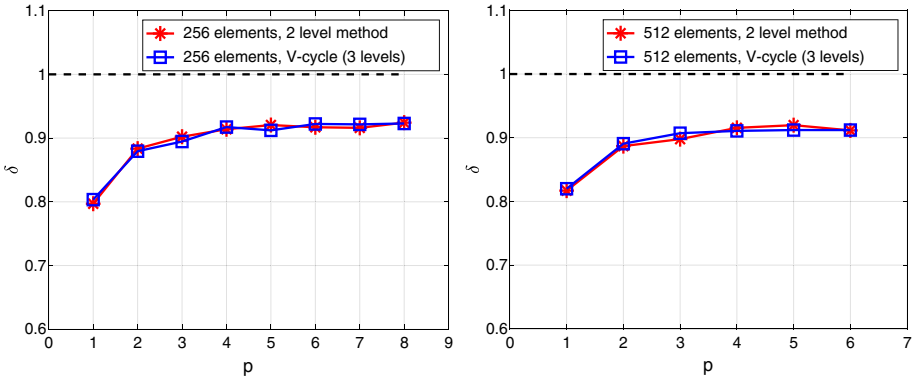

Fig. 3 Estimates of δ_2 and δ_3 in Theorem 6 as a function of p, with $m_1 = m_2 = 3p^2$ and two polygonal grids of 256 (left) and 512 (right) elements

$m_1 = m_2 = m = 3p^2$, cf. Fig. 3. Here, we observe that δ_2 and 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 $m \gtrsim p^2$.

Next, we investigate the performance of the V-cycle algorithm with non-nested partitions presented in Sect. 4. In order to do that, we compute the iteration counts needed by our V-cycle algorithm to reduce the relative residual error below a given tolerance of 10^{-8}, by varying the polynomial degree and the granularity of the finest grid. In Table 1 we report the computed convergence factor

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

where $N_{it,J}$ is the iteration counts needed to reduce the residual below the given tolerance by 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 6, the convergence factor is independent from the spatial discretization step h. Indeed, for a fixed $J \in \{2, 3, 4\}$ and a fixed number of smoothing steps m, the convergence

Table 1 Converge factor ρ_J of the V-cycle multigrid method as a function of m ($C_\sigma^j \equiv C_\sigma = 10$, $p = 1$)

	Set 1			Set 2		
	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			Set 4		
	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 factor ρ_J (and iteration counts) of the V-cycle method as a function of the number m of smoothing steps ($C_\sigma^j \equiv C_\sigma = 10$, $p = 3$)

	Set 1			Set 2		
	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			Set 4		
	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)

factor is roughly constant. In particular, this means that the number of iterations needed by our V-cycle method to attain convergence is independent of the granularity of the underlying grid. As expected, the convergence factor is reduced by increasing the number of smoothing step.

We have repeated the same set of experiments employing $p_j = 3,\ \forall j = 1, \ldots, J$; the results are reported in Table 2 together with the corresponding iteration counts (between parenthesis). First, a comparison between Tables 1 and 2 shows that the convergence factor increases as p grows if the number of smoothing steps is kept fixed. We also observe that, if the number of smoothing step is not chosen sufficiently large, the algorithm might fail to converge. Indeed, according to Theorem 6, in order to attain a uniform convergence (also with respect to p) the number of smoothing steps m must satisfy $m > 2C_A C_Q \gtrsim p^{2+\mu}$, cf. also Fig. 3.

6 Additive Schwarz Smoother

In order to improve the performance of our V-cycle algorithm, we define in this section a domain decomposition preconditioner that can be used as a smoothing operator in place of

of the Richardson one. To this aim, let $\mathcal{T}_j$ and $\mathcal{T}_{j-1}$ be a pair of consecutive (non-nested) coarse/fine meshes, satisfying the grid assumptions given in Sect. 2.1. We next 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 $\kappa_i \in \mathcal{T}_j$:

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

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

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

Here, we define the injection operator from V_j^0 to V_j 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 show 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. The smoothing step is given 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 from the initial guess z_0.

The computed iteration counts based on employing 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, cf. Fig. 1, and we varied the polynomial degree $p \in \{1, 3, 5\}$. First, we observe that, also in this case, the iteration counts seem to be independent of the number of elements in the underlying mesh for a fixed number of smoothing steps m. Moreover, the results show that a minimal number of smoothing steps is not needed to attain the convergence as p increases. Finally,

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 solver 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

Table 4 Iteration counts of the V-cycle solver 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

Table 6 shows the computed convergence factor, where different polynomial approximation degrees are employed on different levels. Also in this case we observe that the iteration counts seem to be independent of the granularity of the underlying grid.

6.1 Applications to Domains with Curved Boundaries

In this section we consider two examples where the coarser grid does not conform to the boundary. Indeed, in these cases the agglomeration process with edge-coarsening might lead to coarse meshes whose boundary do not fit the geometry, cf. Fig. 4 for an example.

In the following we present two examples showing that the convergence properties of our multigrid method seems not to deteriorate for such problems and that our approach seems to

Table 5 Iteration counts of the V-cycle solver 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

Table 6 Iteration counts of the hp-version of the V-cycle solver 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

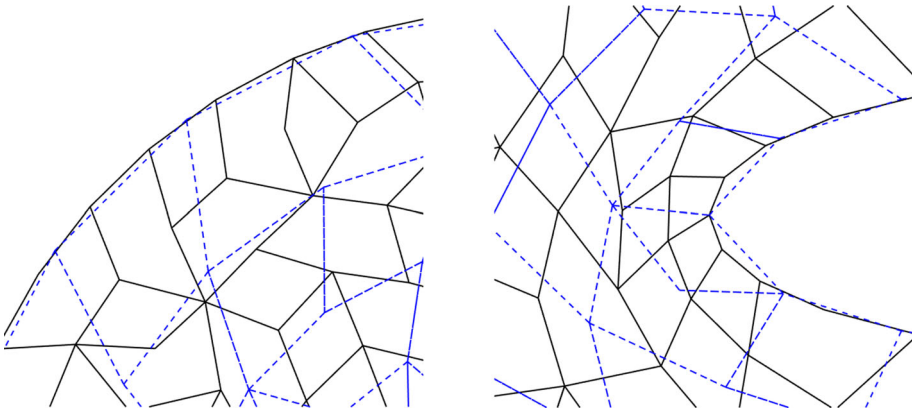

Fig. 4 Examples of fine $\mathcal{T}_h$ (solid line) and coarse $\mathcal{T}_H$ (dashed line) grids for a domain with a curved boundary

be competitive in practical cases. The results of this section have been obtained with the AS smoother, cf. Sect. 6. First, we consider problem (1) with a constant forcing term $f = 1$, and choose the computational domain to be a circular crown $\Omega = \Omega_1 \setminus \Omega_2$, where Ω_1 and Ω_2 are two concentric circles of radius $r_1 = 2$ and $r_2 = \frac{2}{3}$, respectively. We have tested the V-cycle method by defining three sequences of uniform Voronoi grids (set 1, set 2, set 3) as the ones reported in Fig. 5, where, for each set of grids, the first three levels of refinement are shown. Here, each polygonal mesh at different levels is defined independently from the previous one with the only constrain that the cardinality of each coarser grid is approximately $\frac{1}{4}$ of that of the finer level. Tables 7 and 8 show the computed convergence factors for $p = 1$ and $p = 2$, respectively, by choosing $m = 3, 5, 8$ smoothing steps. As expected, the results confirm that the convergence rate depends on p but it is independent of the granularity of the underlying grid as well as the number of levels employed.

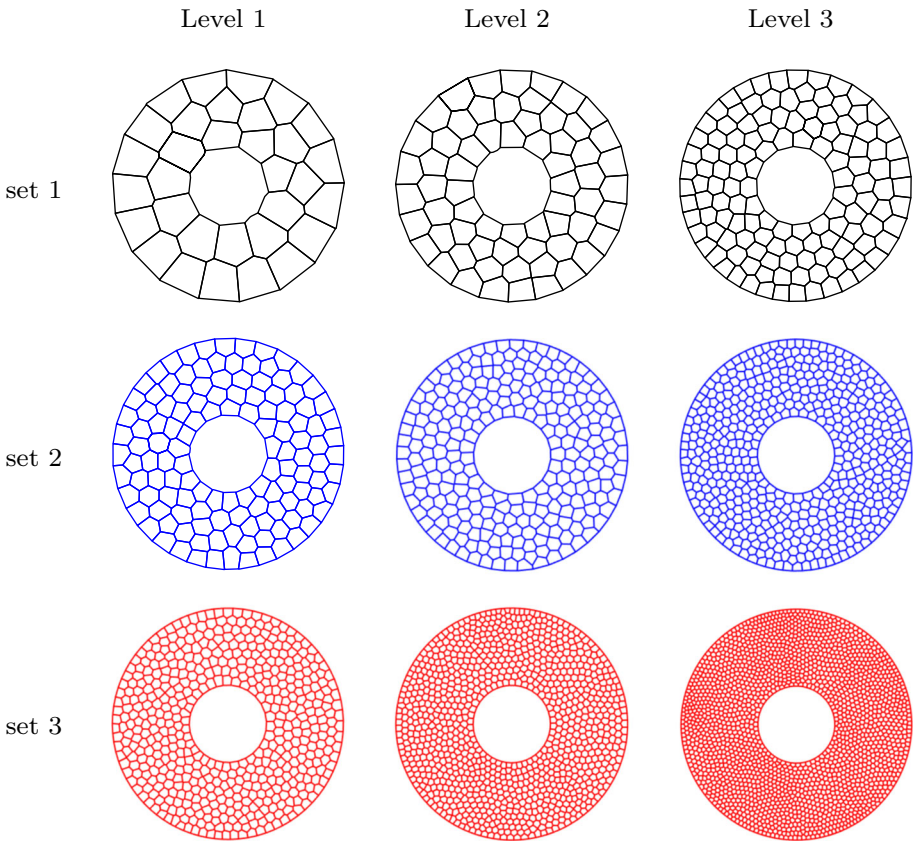

Fig. 5 Circular crown test case: for any set of grids the first three levels of non-nested meshes are shown

Next, we consider the airfoil geometry of [35], which is characterized by a more complicated geometry $\Omega = \Omega_1 \setminus \Omega_2$, where Ω_1 is the circle of radius $r_1 = \frac{3}{2}$, and Ω_2 is the airfoil profile NACA0015 [56]. As before, we consider three sequences of non-nested polygonal meshes (set 1, set 2, set 3), cf. Fig. 6. The grids have been obtained by firstly defining a non-uniform triangular mesh on Ω with the tool *DistMesh* [58], and then by agglomerating based on employing *METIS* [55]. The results for $p = 1$ and $p = 2$ are shown in Tables 9 and 10, respectively. Also in this case we observe that, by fixing the number of smoothing steps m and the polynomial degree p, the convergence factor seems to be independent of the mesh size. Moreover, the performance of the method seems not to deteriorate even if the underlying mesh is characterized by elements of different size, and suggest that our algorithm seems to be well suited for the solution of problems characterized by a local refinement or applications with mesh adaptation.

7 Conclusions

In this paper we have extended the W-cycle multigrid convergence analysis on nested polygonal/polyhedral grids of [8] to V-cycle algorithms with non-nested meshes. We have focused

Table 7 Convergence factor of the h-version of the V-cycle solver with the Additive Schwarz smoother as a function of m (circular crown test case, $p = 1$)

	Set 1			Set 2			Set 3		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	0.268	0.268	0.268	0.274	0.257	0.257	0.325	0.325	0.325
$m = 5$	0.098	0.098	0.098	0.093	0.093	0.093	0.086	0.086	0.086
$m = 8$	0.013	0.013	0.013	0.011	0.011	0.011	0.010	0.010	0.010

Table 8 Convergence factor of the h-version of the V-cycle solver with the Additive Schwarz smoother as a function of m (circular crown test case, $p = 2$)

	Set 1			Set 2			Set 3		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	0.578	0.598	0.598	0.585	0.592	0.592	0.582	0.584	0.583
$m = 5$	0.340	0.340	0.340	0.362	0.367	0.367	0.325	0.332	0.332
$m = 8$	0.105	0.105	0.105	0.125	0.125	0.125	0.121	0.121	0.121

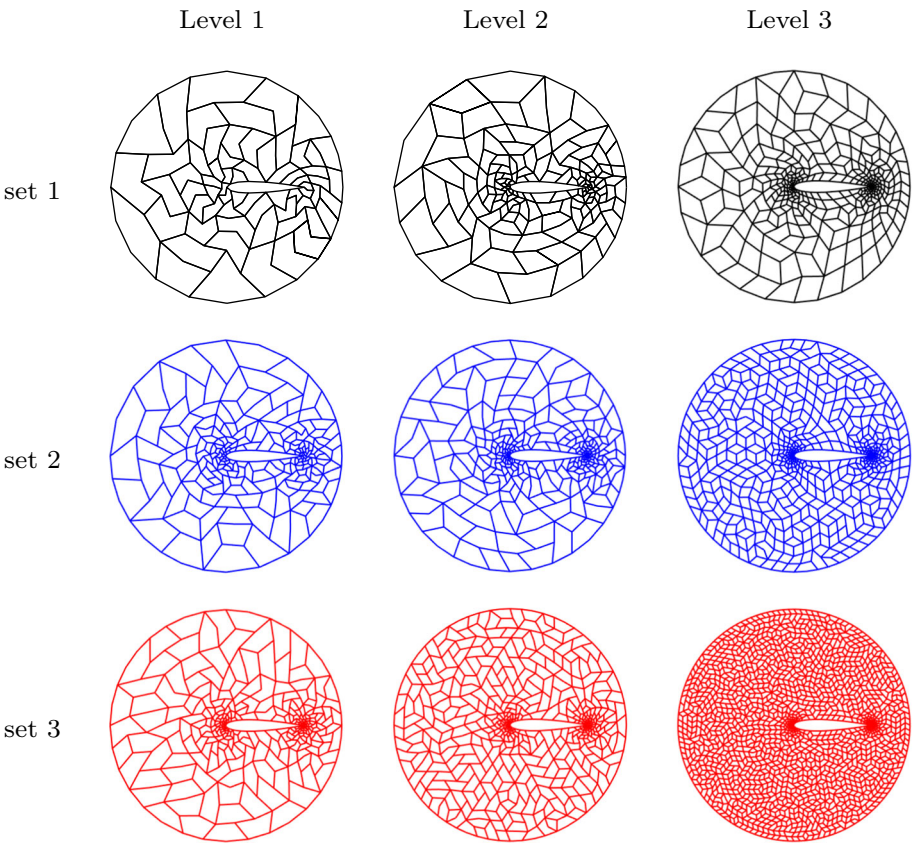

Fig. 6 Airfoil profile test case: for any set of grids the first three levels of non-nested grids are shown

Table 9 Convergence factor of the h-version of the V-cycle solver with the Additive Schwarz smoother as a function of m (airfoil profile test case, $p = 1$)

	Set 1			Set 2			Set 3		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	0.312	0.318	0.318	0.325	0.315	0.315	0.320	0.334	0.334
$m = 5$	0.121	0.124	0.124	0.105	0.107	0.107	0.115	0.124	0.124
$m = 8$	0.020	0.020	0.020	0.022	0.022	0.022	0.021	0.021	0.021

Table 10 Convergence factor of the h-version of the V-cycle solver with the Additive Schwarz smoother as a function of m (airfoil profile test case, $p = 2$)

	Set 1			Set 2			Set 3		
	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.	2 lev.	3 lev.	4 lev.
$m = 3$	0.866	0.848	0.848	0.842	0.848	0.843	0.865	0.864	0.865
$m = 5$	0.621	0.630	0.630	0.629	0.636	0.637	0.655	0.661	0.660
$m = 8$	0.331	0.332	0.334	0.353	0.354	0.355	0.374	0.374	0.376

on the solution of the linear systems of equations stemming from high-order discontinuous Galerkin discretizations of second-order elliptic partial differential equations on polytopic meshes. Here, the possibility of employing non-nested polytopic meshes allows to choose the sequence of grids standing at the basis of the multigrid method based on employing agglomeration procedures together with edge-coarsening. The key aspect of our method is the *projection operator* which is defined as the L^2-projection between two consecutive (non-nested) partitions. By following the general framework introduced in [26] for non-nested multigrid methods, we have proved that our non-nested multigrid method converges uniformly with respect of the number of degree of freedom and the number of multigrid levels, provided that the number of smoothing steps is chosen sufficiently large. More precisely we have proved that the convergence rate is independent of the granularity of the underlying (fine) grid, the polynomial approximation degree p and the number of levels, provided that the number of smoothing steps is chosen of order $p^{2+\mu}$, $\mu \in \{0, 1\}$. We have also proposed a further improvement of the method by considering a Schwarz-type smoother. We demonstrated through several numerical experiments the effectiveness of the proposed algorithm, also for geometries with curved boundaries, where the coarser grid does not fit the geometry. From the implementation point of view, we point out that the assembly of the prolongation and projection matrices needs the knowledge of the intersections between elements of two consecutive levels. Our computations make use of the tool *PolygonClipper* [53], but its extension to the three dimensional case could be expensive. In three dimensions, agglomeration-based procedures which make use of edge-coarsening techniques can also be used to generate the sequence of meshes in the three dimensional case.

Acknowledgements The authors are grateful to the anonymous Reviewers for their valuable comments and constructive suggestions.

Appendix: Proof of Lemma 7

In order to show Lemma 7, we follow the analysis presented in [43]. We first show two preliminary results making use of the properties of 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*

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

Proof Using the triangular inequality, Remark 7 and the approximation estimates of Lemma 4 we have:

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

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

Lemma 9 *Let Assumptions 1–4 hold. Let $g \in L^2(\Omega)$ 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:*

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

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 Corollary 1, we have

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

Using the triangular inequality and Remark 7 we have:

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

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

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

From the elliptic regularity assumption (2) and hypotheses (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:

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

Next, consider the solution z_j of the following problems

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

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 (26). □

References

1. Antonietti, P.F., Ayuso de Dios, B.: Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. Math. Model. Numer. Anal. **41**(1), 21–54 (2007)
2. Antonietti, P.F., Ayuso de Dios, B.: Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems. Math. Model. Numer. Anal. **42**(3), 443–469 (2008)
3. Antonietti, P.F., Brezzi, F., Marini, L.D.: Bubble stabilization of discontinuous Galerkin methods. Comput. Methods Appl. Mech. Eng. **198**(21–26), 1651–1659 (2009)
4. Antonietti, P.F., Facciolà, C., Russo, A., Verani, M.: Discontinuous Galerkin approximation of flows in fractured porous media. MOX report 55/2016 (2016, Submitted)
5. 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)
6. 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)
7. 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)
8. Antonietti, P.F., Houston, P., Hu, X., Sarti, M., Verani, M.: Multigrid algorithms for *hp*-version interior penaty discontinuous Galerkin methods on polygonal and polyhedral meshes. Calcolo **54**(4), 1169–1198 (2017)
9. Antonietti, P.F., Houston, P., Smears, I.: A note on optimal spectral bounds for nonoverlapping domain decomposition preconditioners for *hp*-version discontinuous Galerkin methods. Int. J. Numer. Methods Eng. **13**(4), 513–524 (2016)
10. Antonietti, P.F., Mazzieri, I.: DG methods for the elastodynamics equations on polygonal/polyhedral grids. MOX report 06/2018 (2018, Submitted)
11. 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)
12. Arnold, D.N.: An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Anal. **19**(4), 742–760 (1982)
13. Arnold, D.N., 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)
14. Ayuso de Dios, B., Zikatanov, L.: Uniformly convergent iterative methods for discontinuous Galerkin discretizations. J. Sci. Comput. **40**(1–3), 4–36 (2009)
15. Baker, G.A.: Finite element methods for elliptic equations using nonconforming elements. Math. Comput. **31**(137), 45–59 (1977)
16. Bank, R.E., Dupont, T.: An optimal order process for solving finite element equations. Math. Comput. **36**(153), 35–51 (1981)
17. 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)
18. 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)
19. 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)
20. Braess, D., Verfürth, R.: Multigrid methods for nonconforming finite element methods. SIAM J. Numer. Anal. **22**(4), 979–986 (1990)
21. Bramble, J.H.: Multigrid Methods (Pitman Research Notes in Mathematics Series). Longman Scientific and Technical, New York (1993)
22. 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)
23. Bramble, J.H., Pasciak, J.E.: The analysis of smoothers for multigrid algorithms. Math. Comput. **58**(198), 467–488 (1992)
24. Bramble, J.H., Pasciak, J.E.: New estimates for multilevel algorithms including the V-cycle. Math. Comput. **60**(202), 447–471 (1993)
25. 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)
26. Bramble, J.H., Pasciak, J.E., Xu, J.: The analysis of multigrid algorithms with nonnested space or noninherited quadratic forms. Math. Comput. **56**(193), 1–34 (1991)
27. Bramble, J.H., Zhang, X.: Uniform convergence of the multigrid V-cycle for an anisotropic problem. Math. Comput. **70**(234), 453–470 (2001)

28. Brenner, S.C., Cui, J., Gudi, T., Sung, L.-Y.: Multigrid algorithms for symmetric discontinuous Galerkin methods on graded meshes. Numer. Math. **119**(1), 21–47 (2011)
29. Brenner, S.C., Cui, J., Sung, L.-Y.: Multigrid methods for the symmetric interior penalty method on graded meshes. Numer. Linear Algebra Appl. **16**(6), 481–501 (2009)
30. Brenner, S.C., Owens, L.: A W-cycle algorithm for a weakly over-penalized interior penalty method. Comput. Methods Appl. Mech. Eng. **196**(37–40), 3823–3832 (2007)
31. Brenner, S.C., Park, E.-H., Sung, L.-Y.: A balancing domain decomposition by constraints preconditioner for a weakly over-penalized symmetric interior penalty method. Numer. Linear Algebra Appl. **20**(3), 472–491 (2013)
32. 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)
33. 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)
34. 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)
35. Chan, T.F., Xu, J., Zikatanov, L.: An agglomeration multigrid method for unstructured grids. Contemp. Math. **218**, 67–81 (1998)
36. Cockburn, B., Karniadakis, G.E., Shu, C.-W.: Discontinuous Galerkin Methods. Theory, Computation and Applications. Springer, Berlin (2000)
37. Di Pietro, D.A., Ern, A.: Mathematical Aspects of Discontinuous Galerkin Methods. Springer, Berlin (2011)
38. Dobrev, V.A., Lazarov, R.D., Vassilevski, P.S., Zikatanov, L.: Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. Numer. Linear Algebra Appl. **13**(9), 753–770 (2006)
39. Dryja, M.: On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. Comput. Methods Appl. Mat. **3**(1), 76–85 (2003)
40. Dryja, M., Galvis, J., Sarkis, M.: BDDC methods for discontinuous Galerkin discretization of elliptic problems. J. Complex. **23**, 715–739 (2007)
41. Dryja, M., Krzyżanowski, P., Sarkis, M.: Additive Schwarz method for dG discretization of anisotropic elliptic problems. Lect. Notes Comput. Sci. Eng. **98**, 407–415 (2014)
42. Dryja, M., Sarkis, M.: Additive average Schwarz methods for discretization of elliptic problems with highly discontinuous coefficients. Comput. Methods Appl. Math. **10**(2), 164–176 (2010)
43. Duan, H.Y., Gao, S.Q., Tan, R.C.E., Zhang, S.: A generalized BPX multigrid framework covering nonnested V-cycle methods. Math. Comput. **76**(257), 137–152 (2007)
44. Feng, X., Karakashian, O.A.: Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. SIAM J. Numer. Anal. **39**(4), 1343–1365 (2001). (**electronic**)
45. Feng, X., Karakashian, O.A.: Analysis of two-level overlapping additive Schwarz preconditioners for a discontinuous Galerkin method. In: Domain Decomposition Methods in Science and Engineering (Lyon, 2000), Theory Eng. Appl. Comput. Methods, pp. 237–245. Internat. Center Numer. Methods Eng. (CIMNE), Barcelona (2002)
46. Feng, X., Karakashian, O.A.: Two-level non-overlapping Schwarz preconditioners for a discontinuous Galerkin approximation of the biharmonic equation. J. Sci. Comput. **22**(23), 289–314 (2005)
47. 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)
48. 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)
49. Giani, S., Houston, P.: hp-adaptive composite discontinuous Galerkin methods for elliptic problems on complicated domains. Numer. Methods Part. Differ. Equ. **30**(4), 1342–1367 (2014)
50. Gopalakrishnan, J., Kanschat, G.: A multilevel discontinuous Galerkin method. Numer. Math. **95**(3), 527–550 (2003)
51. Gopalakrishnan, J., Pasciak, J.E.: Multigrid for the Mortar finite element method. SIAM J. Numer. Anal. **37**(3), 1029–1052 (2000)
52. Hesthaven, J.S., Warburton, T.: Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications, 1st edn. Springer, Berlin (2008)
53. Holz, S.: Polygon clipper. https://it.mathworks.com/matlabcentral/fileexchange/8818-polygon-clipper
54. Karakashian, O.A., Collins, C.: Two-level additive Schwarz methods for discontinuous Galerkin approximations of second-order elliptic problems. IMA J. Numer. Anal. **37**(4), 1800–1830 (2017)
55. Karypis, G., Kumar, V.: Metis: Unstructured graph partitioning and sparse matrix ordering system, version 4.0. http://www.cs.umn.edu/~metis (2009)

56. Ladson, C.L., Jr., C.W.: Brooks. Development of a computer program to obtain ordinates for NACA 6- and 6A-series airfoils. NASA Technical Memorandum X-3069 (1974)
57. 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)
58. Persson, P.O., Strang, G.: A simple mesh generator in MATLAB. SIAM Rev. **46**(2), 329–345 (2004)
59. Reed, W.H., Hill, T.R.: Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-479, Los Alamos Scientific Laboratory (1973)
60. Rivière, B.: Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. SIAM, Philadelphia (2008)
61. Scott, L.R., Zhang, S.: Higher-dimensional nonnested multigrid methods. Math. Comput. **58**(198), 457–466 (1992)
62. Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton (1970)
63. Talischi, C., Paulino, G.H., Pereira, A., Menezes, I.F.M.: Polymesher: a general-purpose mesh generator for polygonal elements written in MATLAB. Struct. Multidiscip. Optim. **45**(3), 309–328 (2012)
64. Toselli, A., Widlund, O.B.: Domain Decomposition Methods—Algorithms and Theory. Springer Series in Computational Mathematics, vol. 34. Springer, Berlin (2004)
65. Wheeler, M.F.: An elliptic collocation finite element method with interior penalties. SIAM J. Numer. Anal. **15**(1), 152–161 (1978)
66. 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. Eng. **270**, 113–149 (2014)
67. Xu, X., Chen, J.: Multigrid for the mortar element method for $P1$ nonconforming element. Numer. Math. **88**(2), 381–389 (2001)
68. 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)
69. Zhang, S.: Optimal-order nonnested multigrid methods for solving finite element equations I: on quasi-uniform meshes. Math. Comput. **55**(191), 23–36 (1990)
70. Zhang, S., Zhang, Z.: Treatments of discontinuity and bubble functions in the multigrid method. Math. Comput. **66**(219), 1055–1072 (1997)