MULTIGRID ALGORITHMS FOR hp-DISCONTINUOUS GALERKIN DISCRETIZATIONS OF ELLIPTIC PROBLEMS[*]

PAOLA F. ANTONIETTI[†], MARCO SARTI[†], AND MARCO VERANI[†]

Abstract. We present W-cycle h-, p-, and hp-multigrid algorithms for the solution of the linear system of equations arising from a wide class of hp-version discontinuous Galerkin discretizations of elliptic problems. Starting from a classical framework in geometric multigrid analysis, we define a smoothing and an approximation property, which are used to prove uniform convergence of the W-cycle scheme with respect to the discretization parameters and the number of levels, provided the number of smoothing steps is chosen of order $p^{2+\mu}$, where p is the polynomial approximation degree and $\mu = 0, 1$. A discussion on the effects of employing inherited or noninherited sublevel solvers is also presented. Numerical experiments confirm the theoretical results.

Key words. hp-version discontinuous Galerkin, multigrid algorithms, elliptic problems, geometric multigrid

AMS subject classifications. 65N30, 65N55

DOI. 10.1137/130947015

1. Introduction. Discontinuous Galerkin (DG) methods have undergone a huge development in the last three decades mainly because of their flexibility in dealing within the same unified framework with a wide range of equations, handling nonconforming grids and variable polynomial approximation degrees, and imposing weakly boundary conditions. Therefore, the construction of effective solvers such as domain decomposition and multigrid methods has become an active research field. Domain decomposition methods are based on the definition of subproblems on single subdomains, followed by a *coarse* correction, which ensures the scalability of the method. In the framework of domain decomposition algorithms for DG methods, in [30] a Schwarz preconditioner based on overlapping and nonoverlapping partitions of the domain is analyzed. The case of nonoverlapping Schwarz methods with *inexact* local solvers is addressed in a unified framework in [2, 3]. This topic has been further analyzed in [40, 31, 23, 4, 22, 29, 10, 6]. For substructuring-type preconditioners for DG methods, we mention [28, 27], where Neumann-Neumann and Balancing Domain Decomposition with Constraints (BDDC) for Nitsche-type methods are studied. A unified approach for BDDC was recently proposed in [26], while in [17] a preconditioner for an overpenalized DG method is studied. All these contributions focus on the h-version of the DG method; only recently some attention has been devoted to the development of efficient solvers for hp-DG methods. The first contribution in this direction is in [8], where a nonoverlapping Schwarz preconditioner for the hp-version of DG methods is analyzed; cf. also [7] for the extension to domains with complicated geometrical details. In [24, 21] BDDC and multilevel preconditioners for the hp-version of a DG scheme are analyzed, in parallel with conforming methods. Substructuring-type preconditioners for hp-Nitsche type methods have been studied recently in [5]. The issue of preconditioning hybrid DG methods is investigated in [47].

[†]MOX-Laboratory for Modeling and Scientific Computing, Dipartimento di Matematica, Politecnico di Milano, 20133 Milano, Italy (paola.antonietti@polimi.it, marco.sarti@polimi.it, marco.verani@polimi.it).

Here we are interested in multigrid algorithms for hp-version DG methods that exploit the solution of suitable subproblems defined on different levels of discretization. The levels can differ for the mesh size (h-multigrid), the polynomial approximation degree (p-multigrid), or both (hp-multigrid). In the framework of h-multigrid algorithms for DG methods, in [34] a uniform (with respect to the mesh size) multigrid preconditioner is studied. In [37, 38] a Fourier analysis of a multigrid solver for a class of DG discretizations is performed, focusing on the performance of several relaxation methods, while in [51] the analysis concerns convection-diffusion equations in the convection-dominated regime. Other contributions can be found for low-order DG approximations: in [19] it is shown that V-cycle, F-cycle, and W-cycle multigrid algorithms converge uniformly with respect to all grid levels, with further extensions to an overpenalized method in [16] and graded meshes in [15, 14]. To the best of our knowledge, no theoretical results in the framework of p- and hp-DG methods are available, even though p-multigrid solvers are widely used in practical applications; cf. [32, 41, 43, 42, 49, 11], for example.

In this paper, we present W-cycle h-, p-, and hp-multigrid schemes for high-order DG discretizations of a second-order elliptic problem. We consider a wide class of symmetric DG schemes, and, following the theoretical framework for geometric multigrid methods presented in [18, 19, 15], we prove that the W-cycle algorithms converge uniformly with respect to the granularity of the underlying mesh, the polynomial approximation degree p, and the number of levels, provided that the number of smoothing steps is chosen of order $p^{2+\mu}$ with $\mu = 0, 1$. The numerical experiments confirm our theoretical results and show that our multigrid method converges even if the assumption on the minimum number of smoothing steps is not satisfied, but in this case the convergence factor degenerates when p increases. The key points of our analysis are suitable smoothing and approximation properties of the hp-multigrid method. The smoothing scheme is a Richardson iteration, and we exploit the spectral properties of the stiffness operator to obtain the desired estimates. The approximation property is based on the error estimates for hp-DG methods shown [45, 39, 50]. We also discuss in detail the effects of employing inherited or noninherited sublevel solvers. More precisely, we show that the W-cycle algorithms converge uniformly with respect to the number of levels if noninherited sublevel solvers are employed (i.e., the coarse solvers are constructed rediscretizing our original problem at each level), whereas convergence cannot be independent of the number of levels if inherited bilinear forms are considered (i.e., the coarse solvers are the restriction of the stiffness matrix constructed on the finest grid). Those findings are confirmed by numerical experiments.

The rest of the paper is organized as follows. In section 2 we introduce the model problem and its DG discretization and recall some results needed in the forthcoming analysis. In section 3 we introduce W-cycle schemes based on noninherited bilinear forms, while the convergence analysis is performed in section 4. Multigrid algorithms based on employing inherited bilinear forms are discussed in section 5. Numerical experiments are then presented in section 6. In section 7 we draw some conclusions.

2. Model problem and DG discretization. In this section, we introduce the model problem and its discretization by the hp-version of several symmetric DG methods. We describe the DG formulation on different levels of discretization so as to define at this stage a multilevel framework, which will be a key point in the forthcoming multigrid analysis.

Throughout the paper we will use standard notation for Sobolev spaces [1]. We write $x \lesssim y$ and $x \simeq y$ in lieu of $x \le Cy$ and $C_1 y \le x \le C_2 y$, respectively, for positive

constants C, C_1, and C_2 independent of the discretization parameters. When needed, the constants will be written explicitly.

Let $\Omega \in \mathbb{R}^d$, $d = 2, 3$, be a convex polygonal/polyhedral domain and $f \in L^2(\Omega)$ a given function. We consider the weak formulation of the Poisson problem with homogeneous Dirichlet boundary conditions: find $u \in V := H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$(2.1) \qquad \int_\Omega \nabla u \cdot \nabla v \; dx = \int_\Omega fv \; dx \qquad \forall v \in V.$$

We consider a sequence of quasi-uniform partitions $\{\mathcal{T}_k\}_{k=1}^K$ of Ω, each of which consists of shape-regular elements T of diameter h_T, and we set $h_k := \max_{T \in \mathcal{T}_k} h_T$, $k = 1, \ldots, K$. We suppose that each $T \in \mathcal{T}_k$ is the affine image of a reference element $\widehat{T}$, i.e., $T = \mathsf{F}_T(\widehat{T})$, where $\widehat{T}$ can be either the open, unit simplex or the open, unit hypercube in $\mathbb{R}^d$, $d = 2, 3$. To each $\mathcal{T}_k$, $k = 1, \ldots, K$, we associate a finite dimensional discontinuous space V_k defined as

$$V_k := \{v \in L^2(\Omega) : v \circ \mathsf{F}_T \in \mathbb{M}^{p_k}(\widehat{T}) \quad \forall T \in \mathcal{T}_k\}, \quad k = 1, \ldots, K,$$

where $\mathbb{M}^{p_k}(\widehat{T})$ is either the space of polynomials of total degree less than or equal to $p_k \geq 1$, if $\widehat{T}$ is the reference simplex in $\mathbb{R}^d$, or the space of all tensor-product polynomials on $\widehat{T}$ of degree p_k in each coordinate direction, if $\widehat{T}$ is the reference hypercube in $\mathbb{R}^d$. We note that we consider p_k uniform on $\mathcal{T}_k$. By suitably choosing the sequence $\{\mathcal{T}_k, V_k\}_{k=1}^K$, we can obtain h-, p-, and hp-multigrid methods. More precisely, we retrieve the h-multigrid framework if the polynomial degree is kept fixed for any k (i.e., $p_k = p$) and any $\mathcal{T}_k$, $k = 2, \ldots, K$, derives from $K-1$ successive uniform refinements of an initial (coarse) grid $\mathcal{T}_1$ (so that $h_k = h_1 2^{1-k}$); cf. Figure 1(a). On the other hand, in p-multigrid schemes, the mesh is kept fixed for any k, and from one level to another we vary the polynomial approximation degree in such a way that

$$(2.2) \qquad p_{k-1} \leq p_k \lesssim p_{k-1} \qquad \forall k = 2, \ldots, K;$$

cf. Figure 1(b). Combining the two previous strategies, we retrieve the hp-multigrid method; cf. Figure 1(c). We observe that with the above construction the spaces V_k are nested, i.e., $V_1 \subseteq V_2 \subseteq \cdots \subseteq V_K$.

For any $\mathcal{T}_k$, $k = 1, \ldots, K$, we denote by $\mathcal{F}_k^I$, respectively, $\mathcal{F}_k^B$, the set of interior, respectively, boundary, faces of the partition (if $d = 2$ "face" means "edge"), and set $\mathcal{F}_k := \mathcal{F}_k^I \cup \mathcal{F}_k^B$. Here we have adopted the convention that an interior face is the nonempty intersection of the closure of two neighboring elements. For regular enough vector-valued and scalar functions $\boldsymbol{\tau}$ and v, respectively, we define the *jumps* and *weighted averages* (with $\delta \in [0, 1]$) across the face $F \in \mathcal{F}_k^I$ as follows:

$$
\begin{array}{llll}
[\![\boldsymbol{\tau}]\!] & := \boldsymbol{\tau}^+ \cdot \mathbf{n}_{T^+} + \boldsymbol{\tau}^- \cdot \mathbf{n}_{T^-}, & \{\!\{\boldsymbol{\tau}\}\!\}_\delta & := \delta\boldsymbol{\tau}^+ + (1-\delta)\boldsymbol{\tau}^-, & F \in \mathcal{F}_k^I, \\
[\![v]\!] & := v^+ \mathbf{n}_{T^+} + v^- \mathbf{n}_{T^-}, & \{\!\{v\}\!\}_\delta & := \delta v^+ + (1-\delta)v^-, & F \in \mathcal{F}_k^I,
\end{array}
$$

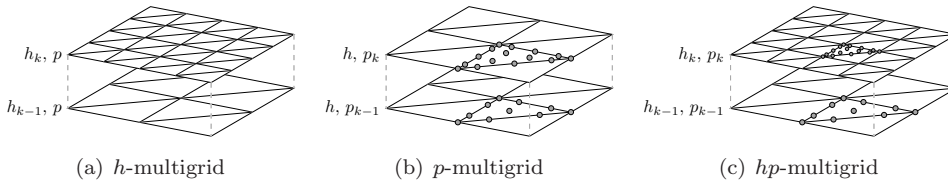|		
(a) h-multigrid	(b) p-multigrid	(c) hp-multigrid

FIG. 1. *Example of the choice of $\{h_k, p_k\}$ for the multigrid methods.*

with $\mathbf{n}_{T^\pm}$ denoting the unit outward normal vector to $\partial T^\pm$ and with $\boldsymbol{\tau}^\pm$ and $v^\pm$ denoting the traces of $\boldsymbol{\tau}$ and v taken within the interior of $T^\pm$, respectively. In the case $\delta = 1/2$ (standard average) we will simply write $\{\!\{\cdot\}\!\}$. On a boundary face $F \in \mathcal{F}_k^B$, we set $[\![v]\!] := v\mathbf{n}_T$, $\{\!\{\boldsymbol{\tau}\}\!\}_\delta := \boldsymbol{\tau}$. We observe that the following relations hold:

$$\{\!\{u\}\!\}_\delta = \{\!\{u\}\!\} + \boldsymbol{\delta} \cdot [\![u]\!], \qquad \{\!\{u\}\!\}_{1-\delta} = \{\!\{u\}\!\} - \boldsymbol{\delta} \cdot [\![u]\!] \quad F \in \mathcal{F}_k^I,$$

with $\boldsymbol{\delta} := (\delta - 1/2)\boldsymbol{n}_F$, where $\boldsymbol{n}_F$ is the outward unit normal vector to the face F to which δ is associated.

We introduce the local lifting operators $r_F^k : [L^2(F)]^d \rightarrow [V_k]^d$ and $l_F^k : L^2(F) \rightarrow [V_k]^d$,

$$\int_\Omega r_F^k(\boldsymbol{\tau}) \cdot \boldsymbol{\eta} \; dx := -\int_F \boldsymbol{\tau} \cdot \{\!\{\boldsymbol{\eta}\}\!\} \; ds \quad \forall \boldsymbol{\eta} \in [V_k]^d \quad \forall F \in \mathcal{F}_k,$$

$$\int_\Omega l_F^k(v) \cdot \boldsymbol{\eta} \; dx := -\int_F v[\![\boldsymbol{\eta}]\!] \; ds \qquad \forall \boldsymbol{\eta} \in [V_k]^d \quad \forall F \in \mathcal{F}_k^I,$$

and set $\mathcal{R}_k(\boldsymbol{\tau}) := \sum_{F \in \mathcal{F}_k} r_F^k(\boldsymbol{\tau})$ and $\mathcal{L}_k(v) := \sum_{F \in \mathcal{F}_k^I} l_F^k(v)$. Next, for any $k = 1, \ldots, K$, we introduce the following bilinear forms $\mathcal{A}_k(\cdot, \cdot) : V_k \times V_k \rightarrow \mathbb{R}$ defined as

$$\mathcal{A}_k(u, v) := \sum_{T \in \mathcal{T}_k} \int_T \nabla u \cdot \nabla v \; dx + \sum_{T \in \mathcal{T}_k} \int_T \nabla u \cdot (\mathcal{R}_k([\![v]\!]) + \mathcal{L}_k(\boldsymbol{\beta} \cdot [\![v]\!])) \; dx$$

$$+ \sum_{T \in \mathcal{T}_k} \int_T (\mathcal{R}_k([\![u]\!]) + \mathcal{L}_k(\boldsymbol{\beta} \cdot [\![u]\!])) \cdot \nabla v \; dx + \mathcal{S}_k^j(u, v)$$

$$+ \theta \sum_{T \in \mathcal{T}_k} \int_T (\mathcal{R}_k([\![u]\!]) + \mathcal{L}_k(\boldsymbol{\beta} \cdot [\![u]\!])) \cdot (\mathcal{R}_k([\![v]\!]) + \mathcal{L}_k(\boldsymbol{\beta} \cdot [\![v]\!])) \; dx,$$

where $\theta = 0$ for the Symmetric Interior Penalty (SIPG) [9] and the weighted Symmetric Interior Penalty (SIPG(δ)) [36] methods and $\theta = 1$ for the Local Discontinuous Galerkin method (LDG) [25]. For the SIPG and SIPG(δ) methods $\boldsymbol{\beta} = \mathbf{0}$ and $\boldsymbol{\beta} = \boldsymbol{\delta}$, respectively, whereas for the LDG method, $\boldsymbol{\beta} \in \mathbb{R}^d$ is a uniformly bounded (and possibly null) vector. The stabilization term $\mathcal{S}_k^j(\cdot, \cdot)$ is defined as

$$\mathcal{S}_k^j(u, v) := \sum_{F \in \mathcal{F}_k} \int_F \sigma_k [\![u]\!] \cdot [\![v]\!] \; ds \quad \forall u, v \in V_k,$$

where $\sigma_k \in L^\infty(\mathcal{F}_k)$ is given by

$$\sigma_k|_F := \frac{\alpha p_k^2}{\min(h_{T^+}, h_{T^-})}, \quad F \in \mathcal{F}_k^I, \sigma_k|_F := \frac{\alpha p_k^2}{h_T}, \quad F \in \mathcal{F}_k^B,$$

with $\alpha \in \mathbb{R}^+$, and $h_{T^\pm}$ diameters of the neighboring elements $T^\pm \in \mathcal{T}_k$.

We are interested in solving efficiently the following DG formulation posed on the finest level K: find $u_K \in V_K$ such that

$$(2.3) \qquad \mathcal{A}_K(u_K, v_K) = \int_\Omega f v_K \; dx \quad \forall v_K \in V_K.$$

Once the basis of V_K has been chosen, we can write (2.3) as the linear system of equations

$$(2.4) \qquad A_K u_K = f_K,$$

where A_K and f_K are the matrix representations of the bilinear form $\mathcal{A}_K(\cdot,\cdot)$ and of the right-hand side of (2.1), respectively, and u_K denotes, with a slight abuse of notation, the vector of the unknowns, given by the coefficients of the expansion of the solution with respect to the basis.

Before describing our multigrid scheme for the solution of (2.4), we recall some technical results that will be needed in the forthcoming analysis.

2.1. Technical results. We endow the space V_k with the DG norm $\|\cdot\|_{\mathrm{DG},k}$ defined as

$$(2.5) \qquad \|v\|_{\mathrm{DG},k}^2 := \sum_{T\in\mathcal{T}_k} \|\nabla v\|_{L^2(T)}^2 + \sum_{F\in\mathcal{F}_k} \|\sigma_k^{1/2}[\![v]\!]\|_{L^2(F)}^2, \ \ k=1,\ldots,K,$$

and observe that, since

$$(2.6) \qquad\qquad h_k \leq h_{k-1} \lesssim h_k \quad \forall k=2,\ldots,K,$$

and thanks to hypothesis (2.2), we have

$$(2.7) \quad \|v_{k-1}\|_{\mathrm{DG},k-1} \leq \|v_{k-1}\|_{\mathrm{DG},k} \lesssim \frac{h_{k-1}^{1/2}}{h_k^{1/2}} \frac{p_k}{p_{k-1}} \|v_{k-1}\|_{\mathrm{DG},k-1} \lesssim \|v_{k-1}\|_{\mathrm{DG},k-1}$$

for any $v \in V_{k-1}$, $k = 2,\ldots,K$. We remark that the hidden constant in the last inequality above depends on the ratio $(h_{k-1}^{1/2}p_k)/(h_k^{1/2}p_{k-1})$, which means that if (2.2) and (2.6) do not hold, such a dependence has to be taken into account.

The following result ensures that the bilinear forms $\mathcal{A}_k(\cdot,\cdot)$ are continuous and coercive in the DG norm (2.5); see [45, Proposition 3.1], [8, Lemma 2.4], or [50, Propositions 2.3 and 2.4] for the proof.

LEMMA 2.1. *For any $k = 1,\ldots,K$, it holds that*

$$(2.8) \qquad\qquad \mathcal{A}_k(u,v) \lesssim \|u\|_{\mathrm{DG},k}\|v\|_{\mathrm{DG},k} \quad \forall u,v \in V_k + V,$$
$$(2.9) \qquad\qquad \mathcal{A}_k(u,u) \gtrsim \|u\|_{\mathrm{DG},k}^2 \qquad\qquad \forall u \in V_k.$$

For the SIPG and SIPG(δ) methods, coercivity holds provided the stabilization parameter α is chosen large enough.

Denoting by $H^s(\mathcal{T}_k)$, $s \geq 1$, the space of elementwise H^s functions, we have the following error estimates, cf. [45, Theorems 3.2 and 3.3], [39, Theorem 4.7], and [50, Theorem 3.3] for the proof.

THEOREM 2.2. *Let $u_k \in V_k$ be the DG solution of problem (2.3) posed on level k, i.e.,*

$$\mathcal{A}_k(u_k, v_k) = \int_\Omega f v_k \ dx \quad \forall v_k \in V_k, \ k = 1,\ldots,K.$$

If the exact solution u of problem (2.1) satisfies $u \in H^{s+1}(\mathcal{T}_k)$, $s \geq 1$, the following error estimates hold:

$$(2.10) \qquad \|u - u_k\|_{\mathrm{DG},k} \lesssim \frac{h_k^{\min(p_k,s)}}{p_k^{s-1/2}} \|u\|_{H^{s+1}(\mathcal{T}_k)}, \qquad k = 1,\ldots,K,$$

$$(2.11) \qquad \|u - u_k\|_{L^2(\Omega)} \lesssim \frac{h_k^{\min(p_k,s)+1}}{p_k^s} \|u\|_{H^{s+1}(\mathcal{T}_k)}, \quad k = 1,\ldots,K.$$

Remark 2.3. We point out that the error estimates of Theorem 2.2 are suboptimal in the polynomial approximation degree p_k of a factor $p_k^{1/2}$ and p_k for (2.10) and (2.11), respectively. Optimal error estimates with respect to p_k can be shown using the projector of [33] provided the solution belongs to a suitable augmented space, or whenever a continuous interpolant can be built; cf. [50]. Therefore, in the following, we will write

$$\|u - u_k\|_{\mathrm{DG},k} \lesssim \frac{h_k^{\min(p_k,s)}}{p_k^{s-\mu/2}} \|u\|_{H^{s+1}(\mathcal{T}_k)}, \qquad k = 1, \dots, K,$$

$$(2.12) \qquad \|u - u_k\|_{L^2(\Omega)} \lesssim \frac{h_k^{\min(p_k,s)+1}}{p_k^{s+1-\mu}} \|u\|_{H^{s+1}(\mathcal{T}_k)}, \quad k = 1, \dots, K,$$

with $\mu = 0, 1$ for optimal and suboptimal estimates, respectively.

Remark 2.4. If we restrict to the case of meshes of d-hypercubes, the forthcoming multigrid analysis can be extended to the methods introduced by Bassi et al. [12] and Brezzi et al. [20], whose bilinear forms can be written as

$$\mathcal{A}_k(u, v) := \sum_{T \in \mathcal{T}_k} \int_T \nabla u \cdot \nabla v \; dx + \sum_{T \in \mathcal{T}_k} \int_T \nabla u \cdot \mathcal{R}_k(\llbracket v \rrbracket) \; dx + \sum_{T \in \mathcal{T}_k} \int_T \mathcal{R}_k(\llbracket u \rrbracket) \cdot \nabla v \; dx$$

$$+ \theta \int_\Omega \mathcal{R}_k(\llbracket u \rrbracket) \cdot \mathcal{R}_k(\llbracket v \rrbracket) dx + \sum_{T \in \mathcal{T}_k} \alpha \int_T r_F^k(\llbracket u \rrbracket) r_F^k(\llbracket v \rrbracket) \; dx,$$

where $\theta = 0, 1$, for the Bassi et al. [12] and Brezzi et al. [20] methods, respectively. Exploiting the equivalence (cf. [48, Lemma 7.2] for the proof)

$$\alpha \|r_F^k(\llbracket v \rrbracket)\|_{L^2(\Omega)}^2 \lesssim \|\sqrt{\sigma_k}\llbracket v \rrbracket\|_{L^2(F)}^2 \lesssim \alpha \|r_F^k(\llbracket v \rrbracket)\|_{L^2(\Omega)}^2, \quad v \in V_k, F \in \mathcal{F}_k,$$

we can prove continuity and coercivity with respect to the DG norm (2.5) using standard techniques. Furthermore, we observe that, in general, the forthcoming analysis holds for any symmetric DG scheme satisfying the continuity and coercivity bounds (2.8) and (2.9), respectively, and the error estimates of Theorem 2.2.

3. Multigrid methods with noninherited sublevel solvers. This section is devoted to the description of the W-cycle algorithm and the proof of few auxiliary results needed in the multigrid analysis. Let n_k be the dimension of V_k. Denoting by $\{\phi_i^k\}_{i=1}^{n_k}$ a set of basis functions of V_k, any $v \in V_k$ can be written as

$$(3.1) \qquad v = \sum_{i=1}^{n_k} v_i \phi_i^k, \quad v_i \in \mathbb{R}, \quad i = 1, \dots, n_k.$$

We will suppose that $\{\phi_i^k\}_{i=1}^{n_k}$ is an orthonormal basis with respect to the $L^2(\widehat{T})$-inner product, $\widehat{T}$ being the reference element. A detailed construction of such a basis can be found in [8]. On V_k we introduce the mesh-dependent inner product

$$(3.2) \qquad (u, v)_k := h_k^d \sum_{i=1}^{n_k} u_i v_i \quad \forall u, v \in V_k, \quad u_i, v_j \in \mathbb{R}, \; i, j = 1, \dots, n_k,$$

where $u_i, v_j \in \mathbb{R}$ are the coefficients of the expansion of u and v, respectively, with respect to the basis $\{\phi_i^k\}_{i=1}^{n_k}$; cf. (3.1).

The next result establishes the connection between the inner product defined in (3.2) and the L^2 norm.

LEMMA 3.1. *For any $v \in V_k$, $k = 1, \ldots, K$, it holds that*

$$\|v\|_{L^2(\Omega)}^2 \simeq (v, v)_k. \tag{3.3}$$

Proof. The proof follows by definition (3.2), the Cauchy–Schwarz inequality, the orthogonality of the basis $\{\phi_i^k\}_{i=1}^{n_k}$, and the fact that $\|\phi_i^k\|_{L^2(T)}^2 \simeq h_k^d$ (cf. [46, Proposition 3.4.1]).   □

We next describe our *hp*-multigrid algorithm for the solution of problem (2.4). To introduce our multigrid method, we need two ingredients (cf. [19, 15, 14]): the intergrid transfer operators (*restriction* and *prolongation*) and a smoothing iteration. The prolongation operator connecting the space V_{k-1} to V_k is denoted by $I_{k-1}^k : V_{k-1} \to V_k$ and consists of the natural injection operator, while the restriction operator $I_k^{k-1} : V_k \to V_{k-1}$ is the adjoint of I_{k-1}^k with respect to the discrete inner product (3.2), i.e.,

$$(v, I_k^{k-1} w)_{k-1} := (I_{k-1}^k v, w)_k \qquad \forall v \in V_{k-1}, w \in V_k.$$

For the smoothing scheme, we choose a Richardson iteration; the corresponding operator is given by

$$B_k := \Lambda_k \mathrm{Id}_k,$$

where Id_k is the identity operator and $\Lambda_k \in \mathbb{R}$ represents a positive bound for the spectral radius of A_k, which is defined as

$$(A_k u, v)_k := \mathcal{A}_k(u, v) \qquad \forall u, v \in V_k, \ k = 1, \ldots, K. \tag{3.4}$$

According to [8, Lemma 2.6] and using the equivalence (3.3), the following estimate for the maximum eigenvalue of A_k can be shown:

$$\lambda_{max}(A_k) \lesssim \frac{p_k^4}{h_k^2}; \tag{3.5}$$

hence,

$$\Lambda_k \lesssim \frac{p_k^4}{h_k^2}. \tag{3.6}$$

As usual, our multigrid algorithm is obtained employing recursion. We then consider the following linear system of equations on level k:

$$A_k z = g$$

with $g \in V_k$. We denote by $\mathsf{MG}_{\mathcal{W}}(k, g, z_0, m_1, m_2)$ the approximate solution obtained by applying the kth level iteration described in Algorithm 1 to the above linear system, with initial guess z_0 and using m_1 pre- and m_2 postsmoothing steps, respectively.

The error propagation operator associated to Algorithm 1 is given by

$$\mathbb{E}_{k,m_1,m_2}(z - z_0) := z - \mathsf{MG}_{\mathcal{W}}(k, g, z_0, m_1, m_2).$$

ALGORITHM 1. W-CYCLE hp-MULTIGRID SCHEME.

---

if k=1 **then**   ▷ Solution on the coarsest level

    Solve $A_k z = g$.

    Set $\mathsf{MG}_{\mathcal{W}}(1, g, z_0, m_1, m_2) = z$.

else

    for $i = 1, \ldots, m_1$ **do**   ▷ Presmoothing

        $z^{(i)} = z^{(i-1)} + B_k^{-1}(g - A_k z^{(i-1)})$;

    end for

    $r_{k-1} = I_k^{k-1}(g - A_k z^{(m_1)})$ ▷ Restriction of the residual

    $\overline{e}_{k-1} = \mathsf{MG}_{\mathcal{W}}(k - 1, r_{k-1}, 0, m_1, m_2)$;

    $e_{k-1} = \mathsf{MG}_{\mathcal{W}}(k - 1, r_{k-1}, \overline{e}_{k-1}, m_1, m_2)$;

    $z^{(m_1+1)} = z^{(m_1)} + I_{k-1}^k e_{k-1}$;   ▷ Prolongation and correction

    for $i = m_1 + 2, \ldots, m_1 + m_2 + 1$ **do**   ▷ Postsmoothing

        $z^{(i)} = z^{(i-1)} + B_k^{-1}(g - A_k z^{(i-1)})$;

    end for

    Set $\mathsf{MG}_{\mathcal{W}}(k, g, z_0, m_1, m_2) = z^{(m_1+m_2+1)}$;

end if

We recall that, according to [35, 13], the following recursive relation holds:

$$(3.7) \quad \begin{cases} \mathbb{E}_{1,m_1,m_2} v & = 0, \\ \mathbb{E}_{k,m_1,m_2} v & = G_k^{m_2}(\mathrm{Id}_k - I_{k-1}^k(\mathrm{Id}_k - \mathbb{E}_{k-1,m_1,m_2}^2)P_k^{k-1})G_k^{m_1} v, \qquad k > 1, \end{cases}$$

where $G_k := \mathrm{Id}_k - B_k^{-1} A_k$ and $P_k^{k-1} : V_k \to V_{k-1}$ is defined as

$$(3.8) \qquad \mathcal{A}_{k-1}(P_k^{k-1}w, v) := \mathcal{A}_k(w, I_{k-1}^k v) \qquad \forall v \in V_{k-1}, w \in V_k.$$

4. Convergence analysis of the multigrid method. In this section we present the convergence analysis of the W-cycle multigrid algorithm described in Algorithm 1. In particular, we first introduce the smoothing property associated to the Richardson iteration, followed by the approximation property. These results are then combined to provide an estimate of the norm of the error propagation operator of the two-level method, which in turn is fundamental to derive an analogous bound for the multigrid scheme.

For $s \in \mathbb{R}$, we first define the norm $\|\!|\cdot|\!\|_{s,k}$ as

$$\|\!|v|\!\|_{s,k}^2 := (A_k^s v, v)_k, \ v \in V_k, \quad k = 1, \ldots, K,$$

and observe that

$$(4.1) \qquad \|\!|v|\!\|_{1,k}^2 = (A_k v, v)_k = \mathcal{A}_k(v, v), \quad \|\!|v|\!\|_{0,k}^2 = (v, v)_k \quad \forall v \in V_k,$$

and, by virtue of (3.3), it holds that

$$(4.2) \qquad \|v\|_{L^2(\Omega)} \simeq \|\!|v|\!\|_{0,k}.$$

Since the bilinear forms $\mathcal{A}_k(\cdot, \cdot)$ are symmetric, in the following we can exploit the eigenvalue problem associated to A_k

$$(4.3) \qquad A_k \psi_i^k = \lambda_i \psi_i^k,$$

where $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_{n_k}$ represent the eigenvalues of A_k and $\{\psi_i^k\}_{i=1}^{n_k}$ are the associated eigenvectors which form a basis for V_k. We can then write any $v \in V_k$ as

$$
(4.4) \qquad v = \sum_{i=1}^{n_k} v_i \psi_i^k, \qquad v_i \in \mathbb{R}.
$$

Next, we introduce the following generalized Cauchy–Schwarz inequality [18, Lemma 6.2.10].

LEMMA 4.1. *For any $v, w \in V_k$ and $s \in \mathbb{R}$, it holds that*

$$
(4.5) \qquad \mathcal{A}_k(v, w) \le \|v\|_{1+s,k} \|w\|_{1-s,k}.
$$

Proof. Considering the eigenvalue problem (4.3) and relation (4.4), it follows that

$$
A_k v = \sum_{i=1}^{n_k} v_i A_k \psi_i^k = \sum_{i=1}^{n_k} v_i \lambda_i \psi_i^k \qquad \forall v \in V_k.
$$

From the definition (3.4) of A_k and of the inner product (3.2), we have

$$
\mathcal{A}_k(v, w) = (A_k v, w)_k = h_k^d \sum_{i=1}^{n_k} v_i w_i \lambda_i = h_k^d \sum_{i=1}^{n_k} v_i \lambda_i^{\frac{1+s}{2}} w_i \lambda_i^{\frac{1-s}{2}}.
$$

The thesis follows applying the Cauchy–Schwarz inequality

$$
(A_k v, w)_k \le \sqrt{h_k^d \sum_{i=1}^{n_k} v_i^2 \lambda_i^{1+s}} \sqrt{h_k^d \sum_{j=1}^{n_k} w_j^2 \lambda_j^{1-s}} = \|v\|_{1+s,k} \|w\|_{1-s,k}. \qquad \square
$$

Before providing the proof of the smoothing property pertaining the Richardson smoother, we introduce an auxiliary result.

LEMMA 4.2. *Given $m \in \mathbb{N} \setminus \{0\}$, it holds that*

$$
\max_{x \in [0,1]} \{ x^\gamma (1-x)^{2m} \} \lesssim (1+m)^{-\gamma}, \quad \gamma = 1, 2.
$$

Proof. The estimate for $\gamma = 1$ is proved in [35, Corollary 6.2.2]. Its extension for $\gamma = 2$ follows by noting that

$$
\max_{x \in [0,1]} \{ x^\gamma (1-x)^{2m} \} = \frac{\gamma^\gamma}{(\gamma + 2m)^{\gamma-1}} \frac{2m^{2m}}{(\gamma + 2m)^{2m+1}}
$$

and employing the estimate for $\gamma = 1$. $\quad \square$

We are now ready to prove the *smoothing property*.

LEMMA 4.3 (smoothing property). *For any $v \in V_k$, it holds that*

$$
(4.6) \qquad \begin{aligned} \|G_k^m v\|_{1,k} &\le \|v\|_{1,k}, \\ \|G_k^m v\|_{s,k} &\lesssim p_k^{2(s-t)} h_k^{t-s} (1+m)^{(t-s)/2} \|v\|_{t,k} \end{aligned}
$$

for $0 \le t < s \le 2$ and $m \in \mathbb{N} \setminus \{0\}$.

Proof. We refer again to the eigenvalue problem (4.3) and write v according to (4.4):

$$
G_k^m v = \left(\text{Id}_k - \frac{1}{\Lambda_k} A_k \right)^m v = \sum_{i=1}^{n_k} \left(1 - \frac{\lambda_i}{\Lambda_k} \right)^m v_i \psi_i^k.
$$

From the above identity we can write

$$\|G_k^m v\|_{s,k}^2 = h_k^d \sum_{i=1}^{n_k} \left(1 - \frac{\lambda_i}{\Lambda_k}\right)^{2m} v_i^2 \lambda_i^s,$$

hence the result for $s = t = 1$ trivially holds. For $t < s$, we have

$$\|G_k^m v\|_{s,k}^2 = \Lambda_k^{s-t} \left\{ h_k^d \sum_{i=1}^{n_k} \left(1 - \frac{\lambda_i}{\Lambda_k}\right)^{2m} \frac{\lambda_i^{s-t}}{\Lambda_k^{s-t}} v_i^2 \lambda_i^t \right\}$$

$$\le \Lambda_k^{s-t} \max_{x \in [0,1]} \{ x^{s-t} (1-x)^{2m} \} \|v\|_{t,k}^2.$$

The thesis follows by estimate (3.6) and Lemma 4.2.   □

Following [19, Lemma 4.2], we now prove the following *approximation property*.

LEMMA 4.4 (approximation property). *Let μ be defined as in Remark* 2.3. *Then,*

$$(4.7) \qquad \|(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})v\|_{0,k} \lesssim \frac{h_{k-1}^2}{p_{k-1}^{2-\mu}} \|v\|_{2,k} \quad \forall v \in V_k.$$

Proof. For any $v \in V_k$, applying (4.2) and the duality formula for the L^2 norm, we obtain

$$(4.8) \qquad \|(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})v\|_{0,k} \lesssim \|(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})v\|_{L^2(\Omega)}$$

$$= \sup_{\substack{\phi \in L^2(\Omega) \\ \phi \neq 0}} \frac{\int_\Omega \phi(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})v \, dx}{\|\phi\|_{L^2(\Omega)}}.$$

Next, for $\phi \in L^2(\Omega)$, let $\eta \in V$ be the solution to

$$\int_\Omega \nabla \eta \cdot \nabla v \, dx = \int_\Omega \phi v \, dx \quad \forall v \in V,$$

and let $\eta_k \in V_k$ and $\eta_{k-1} \in V_{k-1}$ be its DG approximations in V_k and V_{k-1}, respectively, i.e.,

$$(4.9) \quad \mathcal{A}_k(\eta_k, v) = \int_\Omega \phi v \, dx \quad \forall v \in V_k, \quad \mathcal{A}_{k-1}(\eta_{k-1}, v) = \int_\Omega \phi v \, dx \quad \forall v \in V_{k-1}.$$

By (2.12), a standard elliptic regularity result, (2.6), and (2.2), we have

$$(4.10) \qquad \begin{aligned} \|\eta - \eta_k\|_{L^2(\Omega)} &\lesssim \frac{h_k^2}{p_k^{2-\mu}} \|\phi\|_{L^2(\Omega)} \lesssim \frac{h_{k-1}^2}{p_{k-1}^{2-\mu}} \|\phi\|_{L^2(\Omega)}, \\ \|\eta - \eta_{k-1}\|_{L^2(\Omega)} &\lesssim \frac{h_{k-1}^2}{p_{k-1}^{2-\mu}} \|\phi\|_{L^2(\Omega)}. \end{aligned}$$

Moreover, if we consider the definition (3.8) of P_k^{k-1} and (4.9), it holds that

$$\mathcal{A}_{k-1}(P_k^{k-1}\eta_k, w) = \mathcal{A}_k(\eta_k, I_{k-1}^k w) = \mathcal{A}_k(\eta_k, w) = \int_\Omega \phi w \, dx = \mathcal{A}_{k-1}(\eta_{k-1}, w)$$

for any $w \in V_{k-1}$, which implies

$$\eta_{k-1} = P_k^{k-1} \eta_k.$$

Now applying (4.9), the definition (3.8) of P_k^{k-1}, the above identity, the Cauchy–Schwarz inequality (4.5), the L^2 norm equivalence (4.2), and the error estimates (4.10) we obtain

$$
\begin{aligned}
\int_\Omega \phi(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})v \; dx &= \mathcal{A}_k(\eta_k, v) - \mathcal{A}_k(\eta_k, I_{k-1}^k P_k^{k-1}v) \\
&= \mathcal{A}_k(\eta_k, v) - \mathcal{A}_{k-1}(P_k^{k-1}\eta_k, P_k^{k-1}v) \\
&= \mathcal{A}_k(\eta_k, v) - \mathcal{A}_{k-1}(\eta_{k-1}, P_k^{k-1}v) = \mathcal{A}_k(\eta_k - I_{k-1}^k\eta_{k-1}, v) \\
&\leq \|\eta_k - \eta_{k-1}\|_{0,k}\|v\|_{2,k} \lesssim \|\eta_k - \eta_{k-1}\|_{L^2(\Omega)}\|v\|_{2,k} \\
&\leq (\|\eta_k - \eta\|_{L^2(\Omega)} + \|\eta_{k-1} - \eta\|_{L^2(\Omega)})\|v\|_{2,k} \\
&\lesssim \frac{h_{k-1}^2}{p_{k-1}^{2-\mu}}\|\phi\|_{L^2(\Omega)}\|v\|_{2,k}.
\end{aligned}
$$

The above estimate together with (4.8) gives the desired inequality.   □

Lemmas 4.3 and 4.4 allow the convergence analysis of the corresponding two-level method, whose error propagation operator is given by

$$
\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}} = G_k^{m_2}(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})G_k^{m_1}.
$$

THEOREM 4.5.   *There exists a positive constant $C_{2\mathrm{lvl}}$ independent of the mesh size, the polynomial approximation degree, and the level k such that*

$$
(4.11) \qquad \|\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}}v\|_{1,k} \leq C_{2\mathrm{lvl}}\Sigma_k\|v\|_{1,k}
$$

for any $v \in V_k$, $k = 1,\dots,K$, with

$$
\Sigma_k := \frac{p_k^{2+\mu}}{(1+m_1)^{1/2}(1+m_2)^{1/2}},
$$

$m_1, m_2 \geq 1$, *and μ defined as in Remark 2.3. Therefore, the two-level method converges uniformly provided the number of pre- and postsmoothing steps satisfy*

$$
(4.12) \qquad (1+m_1)^{1/2}(1+m_2)^{1/2} \geq \chi p_k^{2+\mu}
$$

for a positive constant $\chi > C_{2\mathrm{lvl}}$.

Proof. Exploiting the smoothing property (4.6), approximation property (4.7), and assumptions (2.6) and (2.2), we obtain

$$
\begin{aligned}
\|\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}}v\|_{1,k} &= \|G_k^{m_2}(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})G_k^{m_1}v\|_{1,k} \\
&\leq Ch_k^{-1}p_k^2(1+m_2)^{-1/2}\|(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})G_k^{m_1}v\|_{0,k} \\
&\leq Ch_k p_k^2 p_{k-1}^{\mu-2}(1+m_2)^{-1/2}\|G_k^{m_1}v\|_{2,k} \\
&\leq C p_k^{2+\mu}(1+m_1)^{-1/2}(1+m_2)^{-1/2}\|v\|_{1,k},
\end{aligned}
$$

and the proof is complete.   □

The next result regards the stability of the intergrid transfer operator I_{k-1}^k and the operator P_k^{k-1}.

LEMMA 4.6.   *There exists a positive constant C_{stab} independent of the mesh size, the polynomial approximation degree, and the level k such that*

$$
(4.13) \qquad \|I_{k-1}^k v\|_{1,k} \leq C_{\mathrm{stab}}\|v\|_{1,k-1} \quad \forall v \in V_{k-1},
$$

$$
(4.14) \qquad \|P_k^{k-1}v\|_{1,k-1} \leq C_{\mathrm{stab}}\|v\|_{1,k} \qquad \forall v \in V_k.
$$

Proof. We apply (4.1), the continuity bound (2.8), and the relation (2.7) between the DG norms on different levels:

$$\|I_{k-1}^k v\|_{1,k}^2 = \mathcal{A}_k(I_{k-1}^k v, I_{k-1}^k v) \lesssim \|I_{k-1}^k v\|_{\mathrm{DG},k}^2 \lesssim \|v\|_{\mathrm{DG},k-1}^2.$$

Estimate (4.13) easily follows using the coercivity (2.9) and denoting by C_{stab}^2 the resulting hidden constant. Inequality (4.14) is obtained by the definition (3.8) of P_k^{k-1}, the Cauchy–Schwarz inequality, (4.1), and (4.13), i.e.,

$$\|P_k^{k-1}v\|_{1,k-1} = \max_{u \in V_{k-1}\setminus\{0\}} \frac{\mathcal{A}_{k-1}(P_k^{k-1}v, u)}{\|u\|_{1,k-1}} = \max_{u \in V_{k-1}\setminus\{0\}} \frac{\mathcal{A}_k(v, I_{k-1}^k u)}{\|u\|_{1,k-1}}$$
$$\leq C_{\mathrm{stab}} \frac{\|v\|_{1,k}\|u\|_{1,k-1}}{\|u\|_{1,k-1}} \leq C_{\mathrm{stab}}\|v\|_{1,k}. \quad \square$$

We are now ready to prove the main result of the paper concerning the convergence of the W-cycle multigrid method.

THEOREM 4.7. *Let Σ_k and $C_{2\mathrm{lvl}}$ be defined as in Theorem 4.5, and let C_{stab} be defined as in Lemma 4.6. Then, there exists a positive constant $\widehat{\mathsf{C}} > C_{2\mathrm{lvl}}$ such that, if the number of pre- and postsmoothing steps satisfies*

$$(4.15) \qquad (1+m_1)^{1/2}(1+m_2)^{1/2} \geq p_k^{2+\mu} \frac{C_{\mathrm{stab}}^2 \widehat{\mathsf{C}}^2}{\widehat{\mathsf{C}} - C_{2\mathrm{lvl}}},$$

it holds that

$$(4.16) \qquad \|\mathbb{E}_{k,m_1,m_2} v\|_{1,k} \leq \widehat{\mathsf{C}}\Sigma_k \|v\|_{1,k} \quad \forall v \in V_k$$

with $\widehat{\mathsf{C}}\Sigma_k < 1$. That is, the W-cycle algorithm converges uniformly with respect to the discretization parameters and the number of levels provided that m_1 and m_2 satisfy (4.15).

Proof. We follow the guidelines given in [15, Theorem 4.6] and proceed by induction. For $k = 1$, (4.16) is trivially true. For $k > 1$ we assume that (4.16) holds for $k-1$. By definition (3.7) we write $\mathbb{E}_{k,m_1,m_2} v$ as

$$\mathbb{E}_{k,m_1,m_2} v = G_k^{m_2}(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1})G_k^{m_1} v + G_k^{m_2} I_{k-1}^k \mathbb{E}_{k-1,m_1,m_2}^2 P_k^{k-1} G_k^{m_1} v,$$

hence

$$\|\mathbb{E}_{k,m_1,m_2} v\|_{1,k} \leq \|\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}} v\|_{1,k} + \|G_k^{m_2} I_{k-1}^k \mathbb{E}_{k-1,m_1,m_2}^2 P_k^{k-1} G_k^{m_1} v\|_{1,k}.$$

The first term can be bounded by Theorem 4.5,

$$\|\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}} v\|_{1,k} \leq C_{2\mathrm{lvl}}\Sigma_k \|v\|_{1,k},$$

while the second term can be estimated by applying the smoothing property (4.6), the stability estimates (4.13)–(4.14), and the induction hypothesis

$$\|G_k^{m_2} I_{k-1}^k \mathbb{E}_{k-1,m_1,m_2}^2 P_k^{k-1} G_k^{m_1} v\|_{1,k} \leq \|I_{k-1}^k \mathbb{E}_{k-1,m_1,m_2}^2 P_k^{k-1} G_k^{m_1} v\|_{1,k}$$
$$\leq C_{\mathrm{stab}} \|\mathbb{E}_{k-1,m_1,m_2}^2 P_k^{k-1} G_k^{m_1} v\|_{1,k}$$
$$\leq C_{\mathrm{stab}} \widehat{\mathsf{C}}^2 \Sigma_{k-1}^2 \|P_k^{k-1} G_k^{m_1} v\|_{1,k}$$
$$\leq C_{\mathrm{stab}}^2 \widehat{\mathsf{C}}^2 \Sigma_{k-1}^2 \|G_k^{m_1} v\|_{1,k}$$
$$\leq C_{\mathrm{stab}}^2 \widehat{\mathsf{C}}^2 \Sigma_{k-1}^2 \|v\|_{1,k}.$$

We then obtain

$$\|\mathbb{E}_{k,m_1,m_2}v\|_{1,k} \leq \left(C_{2\mathrm{lvl}}\Sigma_k + C_{\mathrm{stab}}^2\widehat{\mathsf{C}}^2\Sigma_{k-1}^2\right)\|v\|_{1,k}.$$

By considering the definition of Σ_k given in Theorem 4.5 and (2.2), we obtain

$$\Sigma_{k-1}^2 = \frac{p_{k-1}^{4+2\mu}}{(1+m_1)(1+m_2)} \leq \frac{p_k^{4+2\mu}}{(1+m_1)(1+m_2)} \leq \frac{p_k^{2+\mu}}{(1+m_1)^{1/2}(1+m_2)^{1/2}}\Sigma_k.$$

Therefore,

$$C_{2\mathrm{lvl}}\Sigma_k + C_{\mathrm{stab}}^2\widehat{\mathsf{C}}^2\Sigma_{k-1}^2 \leq \left(C_{2\mathrm{lvl}} + C_{\mathrm{stab}}^2\widehat{\mathsf{C}}^2\frac{p_k^{2+\mu}}{(1+m_1)^{1/2}(1+m_2)^{1/2}}\right)\Sigma_k.$$

We now observe that if m_1 and m_2 satisfy

$$(1+m_1)^{1/2}(1+m_2)^{1/2} \geq p_k^{2+\mu}\frac{C_{\mathrm{stab}}^2\widehat{\mathsf{C}}^2}{\widehat{\mathsf{C}} - C_{2\mathrm{lvl}}},$$

we obtain $C_{2\mathrm{lvl}}\Sigma_k + C_{\mathrm{stab}}^2\widehat{\mathsf{C}}^2\Sigma_{k-1}^2 \leq \widehat{\mathsf{C}}\Sigma_k < 1$, and inequality (4.16) follows. $\square$

5. W-cycle algorithms with inherited bilinear forms. In section 4 we have followed the classical approach in the framework of multigrid algorithms for DG methods [34, 19, 15, 14], where the bilinear forms are assembled on each sublevel. We now consider inherited bilinear forms, that is, the sublevel solvers $\mathcal{A}_k^{\mathrm{R}}(\cdot, \cdot)$ are obtained as the restriction of the original bilinear form $\mathcal{A}_K(\cdot, \cdot)$:

$$(5.1) \qquad \mathcal{A}_k^{\mathrm{R}}(v, w) := \mathcal{A}_K(I_k^K v, I_k^K w) \quad \forall v, w \in V_k \quad \forall k = 1, 2, \ldots, K-1.$$

For $k = 1, \ldots, K-1$, the prolongation operators are defined as $I_k^K := I_{K-1}^K I_{K-2}^{K-1} \cdots I_k^{k+1}$, where I_k^{k+1} is defined as before. The associated operators A_k^{R} can be computed as $A_k^{\mathrm{R}} = I_K^k A_K I_k^K$. Using the new definition of the sublevel solvers, it is easy to see that the coercivity estimate remains unchanged, i.e., $\mathcal{A}_k^{\mathrm{R}}(u, u) \gtrsim \|u\|_{\mathrm{DG},k}^2$ for all $u \in V_k$, whereas the continuity bound (2.8) modifies as follows:

$$(5.2) \quad \mathcal{A}_k^{\mathrm{R}}(u, v) \lesssim \|u\|_{\mathrm{DG},K}\|v\|_{\mathrm{DG},K} \lesssim \frac{p_K^2}{p_k^2}\frac{h_k}{h_K}\|u\|_{\mathrm{DG},k}\|v\|_{\mathrm{DG},k} \quad \forall u, v \in V(h_k).$$

The modified continuity bound affects the error estimates and it can be proved that the error estimate in the L^2 norm on level V_k (cf. (2.12)) becomes

$$(5.3) \qquad \|u - u_k\|_{L^2(\Omega)} \lesssim \frac{p_K^4}{p_k^4}\frac{h_k^2}{h_K^2}\frac{h_k^{\min(p_k, s)+1}}{p_k^s}\|u\|_{H^{s+1}(\mathcal{T}_k)}.$$

Moreover, also estimate (3.5) is affected by (5.2), and it becomes

$$(5.4) \qquad\qquad \lambda_{max}(A_k^{\mathrm{R}}) \lesssim \frac{p_K^2 p_k^2}{h_K h_k},$$

where we have used the continuity bound (5.2) and the inverse and trace inequalities as in [8]. We then consider the Richardson smoothing scheme with
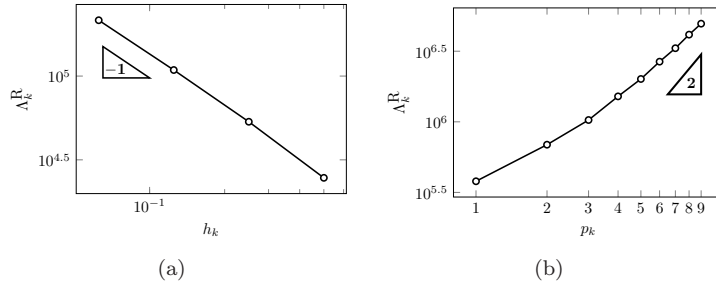
$$B_k^{\mathrm{R}} = \Lambda_k^{\mathrm{R}}\mathrm{Id}_k,$$

FIG. 2. *Values of Λ_k^{R} as a function of* (a) *the mesh size h_k and* (b) *the polynomial approximation degree p_k.*

where, by (5.4), $\Lambda_k^{\mathrm{R}} \in \mathbb{R}$ is such that

$$
(5.5) \qquad \Lambda_k^{\mathrm{R}} \lesssim \frac{p_K^2 p_k^2}{h_K h_k}.
$$

In Figure 2, we verify estimate (5.5); in particular, in Figure 2(a) we keep the finest grid fixed and $p_k = 2$ for any k and compute Λ_k^{R} for A_k^{R} obtained as the restriction on sublevels of decreasing mesh size, while in Figure 2(b), starting from a finest level corresponding to $p_K = 10$, the operators A_k^{R} are computed as restrictions on sublevels with fixed mesh size and such that $p_k = p_{k+1} - 1$.

The W-cycle algorithm is defined as in Algorithm 1 but replacing A_k and B_k with A_k^{R} and B_k^{R}, respectively. The convergence analysis then follows the lines of section 4 and we prove that a dependence on the number of levels is introduced. The following modified smoothing property can be proved by reasoning as in the proof of Lemma 4.3.

LEMMA 5.1. *For any $v \in V_k$, it holds that*

$$
(5.6) \qquad
\begin{aligned}
&\|G_k^m v\|_{1,k} \lesssim \|v\|_{1,k}, \\
&\|G_k^m v\|_{s,k} \lesssim (p_K p_k)^{(s-t)} (h_K h_k)^{(t-s)/2} (1+m)^{(t-s)/2} \|v\|_{t,k}
\end{aligned}
$$

with $0 \leq t < s \leq 2$.

From the error estimate (5.3), a different approximation property follows, which is reported in the following result.

LEMMA 5.2 (approximation property). *Let μ be defined as in Remark* 2.3. *Then,*

$$
(5.7) \qquad \|(\mathrm{Id}_k - I_{k-1}^k P_k^{k-1}) v\|_{0,k} \lesssim \frac{p_K^4}{p_{k-1}^4} \frac{h_{k-1}^2}{h_K^2} \frac{h_{k-1}^2}{p_{k-1}^{2-\mu}} \|v\|_{2,k} \quad \forall v \in V_k.
$$

Regarding the convergence of the two-level method, estimates (5.6) and (5.7) introduce a dependence on the level k, as shown in the next result.

THEOREM 5.3. *There exists a positive constant $C_{2\mathrm{lvl}}^{\mathrm{R}}$ independent of the mesh size, the polynomial approximation degree, and the level k such that*

$$
\|\mathbb{E}_{k,m_1,m_2}^{2\mathrm{lvl}} v\|_{1,k} \leq C_{2\mathrm{lvl}}^{\mathrm{R}} \Sigma_k^{\mathrm{R}} \|v\|_{1,k}
$$

for any $v \in V_k$, with

$$
(5.8) \qquad \Sigma_k^{\mathrm{R}} := 2^{3(K-k)} \frac{p_K^6 p_k^{\mu-4}}{(1+m_1)^{1/2}(1+m_2)^{1/2}},
$$

and μ defined as in Remark 2.3.

We observe that the term $2^{3(K-k)}$ in (5.8) is due to the fact that $h_k = 2^{K-k} h_K$. We also note that, from definition (5.1), the stability estimates (4.13) and (4.14) reduce to

$$\|I_{k-1}^k v\|_{1,k} = \|v\|_{1,k-1} \quad \forall v \in V_{k-1}, \qquad \|P_k^{k-1} v\|_{1,k-1} \le \|v\|_{1,k} \quad \forall v \in V_k,$$

thus resulting in the following convergence estimate for the W-cycle algorithm.

THEOREM 5.4. *Let Σ_k^R and C_{2lvl}^R be defined as in Theorem 5.3. Then, there exists a positive constant $\mathsf{C}^R > C_{2lvl}^R$ such that if the number of pre- and postsmoothing steps satisfies*

$$(1+m_1)^{-1/2}(1+m_2)^{-1/2} \ge 2^{3(K-k+2)} p_K^6 p_{k-1}^{2\mu-8} p_k^{4-\mu} \frac{(\mathsf{C}^R)^2}{\mathsf{C}^R - C_{2lvl}^R},$$

it holds that

$$\|\mathbb{E}_{k,m_1,m_2} v\|_{1,k} \le \mathsf{C}^R \Sigma_k^R \|v\|_{1,k} \quad \forall v \in V_k$$

with $\mathsf{C}^R \Sigma_k^R < 1$ and μ defined as in Remark 2.3.

6. Numerical results. In this section we report some numerical experiments to verify the sharpness of our theoretical estimates and to test the practical performance of our W-cycle algorithms.

We first verify numerically the smoothing (Lemma 4.3) and approximation (Lemma 4.4) properties of the h-multigrid scheme. Since the dependence on the mesh size is well known, we restrict ourselves to investigating the dependence on the polynomial approximation degree and on the number of smoothing steps. To this aim, we consider a Cartesian grid $\mathcal{T}_K$ of the unit square $\Omega = (0,1)^2$ with $h_K = 0.125$ and the SIPG and LDG discretizations with $\alpha = 10$. In Figure 3(a), we report the estimate of the smoothing property constant, $s = 2$ and $t = 0$ in (4.6), as a function of $p_K = p = 1, 2, \ldots, 10$, employing $m = 2$ smoothing steps. We observe that the numerical data underpin the theoretical result and that the expected rates seem to be achieved asymptotically. We have repeated the same set of experiments by fixing the polynomial order $p_k = 2$ and varying the number of smoothing steps m; the results are reported in Figure 3(b). Also in this case the expected rates seem to be achieved asymptotically. To verify the approximation property, we consider a coarse mesh $\mathcal{T}_{K-1}$ of size $h_{K-1} = 0.25$ such that $\mathcal{T}_K$ is obtained by a uniform refinement of $\mathcal{T}_{K-1}$, and we set $p_K = p_{K-1} = p$. The bound of Lemma 4.4 is verified in Figure 3(c), where the approximation property constant as a function of $p = 1, 2, \ldots, 10$ is shown. The same behavior has been observed fixing the mesh size and varying the polynomial degree between level K and $K - 1$; for brevity these results have been omitted. In Figure 3(d), we investigate the behavior of the constants $C_{2lvl}\Sigma_k$ and $\widehat{C}\Sigma_k$ appearing in the estimates (4.11) and (4.16), respectively, as a function of the polynomial approximation degree. To this aim we have considered the h-multigrid scheme with $K = 2, 3$ (i.e., two and three levels), with $m_1 = m_2 = m = 3p^2$ smoothing steps and $p_k = p$, $k = 1, \ldots, K$. As predicted in Theorems 4.5 and 4.7, these quantities are constant (and strictly lower than one) as p grows, provided that the number of smoothing steps is chosen of order p^2.

Next, we investigate the performance of the W-cycle multigrid scheme described in Algorithm 1 in terms of the convergence factor,

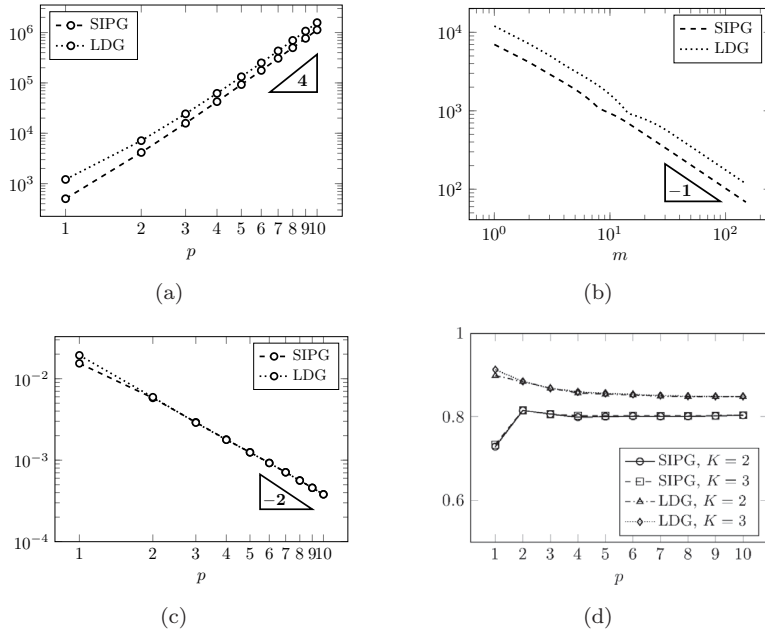$$\rho := \exp\left(\frac{1}{N} \ln \frac{\|\mathbf{r}_N\|_2}{\|\mathbf{r}_0\|_2}\right),$$

(a)

(b)

(c)

(d)

FIG. 3. *Estimates of the smoothing property constant (Lemma 4.3) as a function of* (a) p
and (b) m. (c) *Estimates of the approximation property constant (Lemma 4.4) as a function of p.*
(d) *Estimates of $C_{2lvl}\Sigma_k$ and $\widehat{C}\Sigma_k$ in (4.11) and (4.16), respectively, as a function of p, for $K = 2, 3$
and $m_1 = m_2 = m = 3p^2$. SIPG and LDG methods ($\alpha = 10$), Cartesian grids.*

where N are the iteration counts needed to achieve convergence up to a (relative)
tolerance of 10^{-8} and $\mathbf{r}_N$ and $\mathbf{r}_0$ are the final and initial residuals, respectively. We
start from the h-multigrid scheme, and for all the test cases we fix the coarsest mesh $\mathcal{T}_1$
of size $h_1 = 0.25$, consisting of triangular/Cartesian elements (cf. Figure 4), and build
a sequence of uniform refined grids. This implies that the mesh size of the resulting
grid $\mathcal{T}_K$ obtained after $K-1$ refinements decreases by increasing the number of levels.
This allows us to verify at the same time the uniformity of the method with respect to
the granularity of the mesh and the number of levels. The polynomial approximation
degree is kept fixed, $p_k = p = 1$, $k = 1, \ldots, K$, and we set the penalization parameter
$\alpha = 10$. Figure 5(a) and Figure 5(b) show the computed convergence factors as a
function of m (with $m_1 = m_2 = m$) and the number of levels, obtained with the SIPG
on Cartesian grids and with the LDG on structured triangular grids, respectively.
As predicted by Theorem 4.7, we observe that the rate of convergence is independent
of the number of levels k, and for the LDG method a minimum number of smoothing
steps is needed, retrieving the well-known results already reported in [19, 14]. For the
sake of completeness, in Figure 5(c) we also test the case of inherited bilinear forms
(cf. Theorem 5.4) obtained by considering A_k^{R} instead of A_k for the SIPG method
on structured triangular grids. We observe that the convergence factor increases with
the number of levels, which means that qualitatively the data support the theoretical
prediction.

We now focus on the dependence of the convergence factor of the h-multigrid
scheme on the polynomial approximation degree, and in Table 1 we report the values
of ρ of the h-multigrid scheme for the SIPG method (structured triangular grids) as a
function of the number of levels and the polynomial approximation degree (we recall
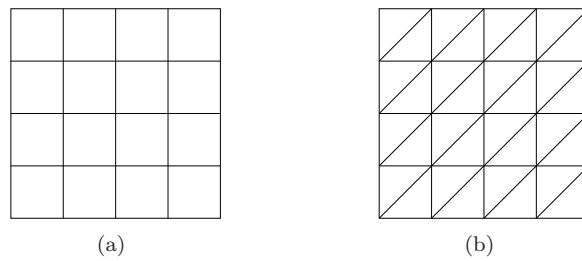that $p_k = p$, $k = 1, \ldots, K$) with $m_1 = m_2 = m = p^2$. As expected, with such a choice

FIG. 4. *Examples of the coarsest grid* $\mathcal{T}_1$: (a) *Cartesian grid and* (b) *structured triangular grid.*
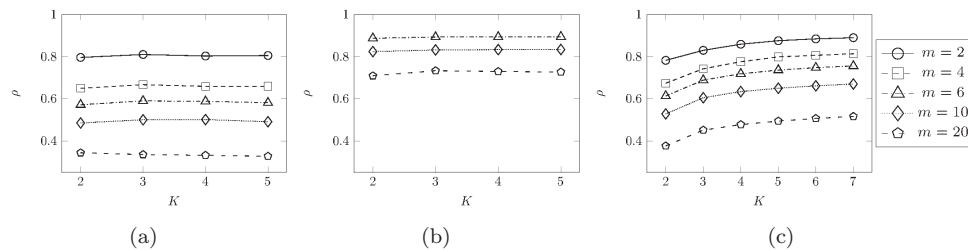

FIG. 5. *Convergence factor* ρ *of the h-multigrid scheme as a function of the number of smoothing steps* $m_1 = m_2 = m$ *and the number of levels K for* (a) *SIPG method on Cartesian grids,* (b) *LDG method on triangular structured grids, and* (c) *SIPG method on structured triangular grids with* A_k^{R} *($\alpha = 10$, $p = 1$).*

of the number of smoothing steps, the convergence factor is substantially constant, thus confirming the conclusions already drawn from the data reported in Figure 3(d).

We next show that our multigrid algorithm indeed converges even if the assumption on the minimum number of smoothing steps is not satisfied, but in this case the performance of the algorithm deteriorates when increasing the polynomial approximation degree. In Table 2, the convergence factor of the h-multigrid is reported as a function of the polynomial approximation degree p and for different levels K, for both the SIPG and LDG methods. Here $m_1 = m_2 = m = 6$ for any p, which implies that the hypotheses (4.12) and (4.15) on the minimum number of smoothing steps to guarantee convergence are not satisfied.

Analogous results have been obtained employing the p-multigrid algorithm. In this test case, we fix the mesh size $h_k = h = 0.0625$ for any $k = 1, \ldots, K$, and we set $p_{k-1} = p_k - 1$ with the convention that $p_K = p$. In Table 3 we report the convergence factor as a function of the number of smoothing steps $m_1 = m_2 = m$ and the number of levels for $p = 5$. Since, with $p_{k-1} = p_k - 1$, the ratio p_k / p_{k-1} is not constant among the levels, uniformity with respect to the number of levels seems to be achieved only asymptotically.

In Table 4, we fix the number of smoothing steps ($m_1 = m_2 = 10$), vary the polynomial approximation degree $p = 2, 3, \ldots, 6$, and report the convergence factor of the p-multigrid method. As before, we address the performance of both the SIPG and LDG methods on Cartesian and structured triangular grids, respectively. The results reported in Table 4 show that if m is constant, whereas the h-version of the method ensures uniformity with respect to the mesh size, the p-multigrid scheme in general does not guarantee independence of the polynomial approximation degree p.

TABLE 1

Convergence factor ρ of the h-multigrid scheme as a function of the number of levels K and the polynomial approximation degree p ($m_1 = m_2 = m = p^2$, $\alpha = 10$). SIPG method on structured triangular grids.

	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$K=2$	0.8717	0.8741	0.8789	0.8821	0.8732	0.8738
$K=3$	0.8777	0.8848	0.8861	0.8675	0.8627	0.8656
$K=4$	0.8791	0.8824	0.8755	0.8529	0.8586	0.8627
$K=5$	0.8798	0.8783	0.8625	0.8467	0.8542	0.8583
	$m=1$	$m=4$	$m=9$	$m=16$	$m=25$	$m=36$

TABLE 2

Convergence factor ρ of the h-multigrid scheme as a function the polynomial approximation degree p and the number of levels K ($m_1 = m_2 = m = 6$, $\alpha = 10$).

	SIPG, Cartesian grids					LDG, triangular grids				
	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$K=2$	0.862	0.904	0.940	0.956	0.967	0.937	0.963	0.976	0.978	0.979
$K=3$	0.861	0.900	0.928	0.949	0.963	0.942	0.965	0.973	0.972	0.972
$K=4$	0.857	0.884	0.927	0.950	0.963	0.941	0.961	0.967	0.969	0.975

TABLE 3

Convergence factor ρ of the p-multigrid scheme as a function of the number of smoothing steps $m_1 = m_2 = m$ and the number of levels K ($p = 5$, $\alpha = 10$).

	SIPG, Cartesian grids					LDG, triangular grids				
	$m=2$	$m=4$	$m=6$	$m=10$	$m=20$	$m=2$	$m=4$	$m=6$	$m=10$	$m=20$
$K=2$	0.931	0.877	0.852	0.818	0.725	0.953	0.919	0.892	0.857	0.791
$K=3$	0.934	0.882	0.854	0.818	0.725	0.956	0.921	0.888	0.854	0.789
$K=4$	0.951	0.919	0.892	0.841	0.734	0.972	0.949	0.930	0.895	0.818

TABLE 4

Convergence factor of the p-multigrid scheme as a function of the polynomial approximation degree p and the number of levels K ($m_1 = m_2 = m = 10$, $\alpha = 10$).

	SIPG, Cartesian grids					LDG, triangular grids				
	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$	$p=2$	$p=3$	$p=4$	$p=5$	$p=6$
$K=2$	0.576	0.788	0.771	0.818	0.849	0.807	0.837	0.851	0.857	0.881
$K=3$	-	0.784	0.779	0.818	0.849	-	0.841	0.856	0.854	0.879
$K=4$	-	-	0.781	0.841	0.844	-	-	0.879	0.895	0.890

7. Conclusions. We have analyzed a W-cycle *hp*-multigrid scheme for high-order DG discretizations of elliptic problems. We have shown uniform convergence with respect to the discretization parameters, provided the number of pre- and postsmoothing steps is chosen of order $p^{2+\mu}$, p being the polynomial approximation degree and $\mu = 0, 1$. Besides the traditional approach, where the coarse matrices are built on each level [34, 19, 15, 14], we have also considered the case of inherited bilinear forms, showing that the rate of convergence cannot be uniform with respect to the number of levels. Finally, the theoretical results obtained in this paper pave the way for future developments in enhancing multigrid methods for high-order DG discretizations by introducing more sophisticated smoothing schemes and whose performance can be compared to that of other solution techniques; see, e.g., the algebraic multigrid preconditioner of [44]. Such an issue will be the subject of future research.

REFERENCES

[1] R. A. ADAMS, *Sobolev spaces*, Pure and Applied Mathematics 65, Academic Press, New York, 1975.

[2] P. F. ANTONIETTI AND B. AYUSO, *Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: Non-overlapping case*, M2AN Math. Model. Numer. Anal., 41 (2007), pp. 21–54.

[3] P. F. ANTONIETTI AND B. AYUSO, *Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems*, M2AN Math. Model. Numer. Anal., 42 (2008), pp. 443–469.

[4] P. F. ANTONIETTI AND B. AYUSO, *Two-level Schwarz preconditioners for super penalty discontinuous Galerkin methods*, Commun. Comput. Phys., 5 (2009), pp. 398–412.

[5] P. F. ANTONIETTI, B. AYUSO, S. BERTOLUZZA, AND M. PENNACCHIO, *Substructuring preconditioners for an hp domain decomposition method with interior penalty mortaring*, Calcolo, (2014).

[6] P. F. ANTONIETTI, B. AYUSO, S. C. BRENNER, AND L.-Y. SUNG, *Schwarz methods for a preconditioned WOPSIP method for elliptic problems*, Comput. Methods Appl. Math., 12 (2012), pp. 241–272.

[7] P. F. ANTONIETTI, S. GIANI, AND P. HOUSTON, *Domain decomposition preconditioners for discontinuous Galerkin methods for elliptic problems on complicated domains*, J. Sci. Comput., 60 (2014), pp. 203–227.

[8] P. F. ANTONIETTI AND P. HOUSTON, *A class of domain decomposition preconditioners for hp-discontinuous Galerkin finite element methods*, J. Sci. Comput., 46 (2011), pp. 124–149.

[9] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[10] A. T. BARKER, S. C. BRENNER, AND L.-Y. SUNG, *Overlapping Schwarz domain decomposition preconditioners for the local discontinuous Galerkin method for elliptic problems*, J. Numer. Math., 19 (2011), pp. 165–187.

[11] F. BASSI, A. GHIDONI, S. REBAY, AND P. TESINI, *High-order accurate p-multigrid discontinuous Galerkin solution of the Euler equations*, Internat. J. Numer. Methods Fluids, 60 (2009), pp. 847–865.

[12] F. BASSI, S. REBAY, G. MARIOTTI, S. PEDINOTTI, AND M. SAVINI, *A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows*, in Proceedings of the 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics, R. Decuypere and G. Dibelius, eds., Technologisch Instituut, Antwerpen, Belgium, 1997, pp. 99–108.

[13] J. BRAMBLE, *Multigrid Methods*, Pitman Res. Notes Math. 294, Longman Scientific and Technical, Harlow, UK, 1993.

[14] S. C. BRENNER, J. CUI, T. GUDI, AND L.-Y. SUNG, *Multigrid algorithms for symmetric discontinuous Galerkin methods on graded meshes*, Numer. Math., 119 (2011), pp. 21–47.

[15] S. C. BRENNER, J. CUI, AND L.-Y. SUNG, *Multigrid methods for the symmetric interior penalty method on graded meshes*, Numer. Linear Algebra Appl., 16 (2009), pp. 481–501.

[16] S. C. BRENNER AND L. OWENS, *A W-cycle algorithm for a weakly over-penalized interior penalty method*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 3823–3832.

[17] S. C. BRENNER, E.-H. PARK, AND L.-Y. SUNG, *A balancing domain decomposition by constraints preconditioner for a weakly over-penalized symmetric interior penalty method*, Numer. Linear Algebra Appl., 20 (2013), pp. 472–491.

[18] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Appl. Math. 15, Springer, New York, 2008.

[19] S. C. BRENNER AND J. ZHAO, *Convergence of multigrid algorithms for interior penalty methods*, Appl. Numer. Anal. Comput. Math., 2 (2005), pp. 3–18.

[20] F. BREZZI, G. MANZINI, D. MARINI, P. PIETRA, AND A. RUSSO, *Discontinuous finite elements for diffusion problems*, in Atti convegno in onore di F. Brioschi, Istituto Lombardo, Accademia di Scienze e Lettere, 1999.

[21] K. BRIX, M. CAMPOS PINTO, C. CANUTO, AND W. DAHMEN, *Multilevel preconditioning of discontinuous Galerkin spectral element methods. Part I: geometrically conforming meshes*, IMA J. Numer. Anal. (2014).

[22] K. BRIX, M. CAMPOS PINTO, W. DAHMEN, AND R. MASSJUNG, *Multilevel preconditioners for the interior penalty discontinuous Galerkin method II: Quantitative studies*, Commun. Comput. Phys., 5 (2009), pp. 296–325.

[23] K. BRIX, M. PINTO, AND W. DAHMEN, *A multilevel preconditioner for the interior penalty discontinuous Galerkin method*, SIAM J. Numer. Anal., 46 (2008), pp. 2742–2768.

[24] C. Canuto, L. F. Pavarino, and A. Pieri, *BDDC preconditioners for continuous and discontinuous Galerkin methods using spectral/hp elements with variable local polynomial degree*, IMA J. Numer. Anal., 34 (2014), pp. 879–903.

[25] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.

[26] L. T. Diosady and D. L. Darmofal, *A unified analysis of balancing domain decomposition by constraints for discontinuous Galerkin discretizations*, SIAM J. Numer. Anal., 50 (2012), pp. 1695–1712.

[27] M. Dryja, J. Galvis, and M. Sarkis, *BDDC methods for discontinuous Galerkin discretization of elliptic problems*, J. Complexity, 23 (2007), pp. 715–739.

[28] M. Dryja and M. Sarkis, *A Neumann-Neumann Method for DG Discretization of Elliptic Problems*, Informes de matemática, Institut. Nacional de Matemática Pura e Aplicada, 2006.

[29] M. Dryja and M. Sarkis, *Additive average Schwarz methods for discretization of elliptic problems with highly discontinuous coefficients*, Comput. Methods Appl. Math., 10 (2010), pp. 164–176.

[30] X. Feng and O. A. Karakashian, *Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems*, SIAM J. Numer. Anal., 39 (2001), pp. 1343–1365.

[31] X. Feng and O. A. Karakashian, *Two-level non-overlapping Schwarz preconditioners for a discontinuous Galerkin approximation of the biharmonic equation*, J. Sci. Comput., 22/23 (2005), pp. 289–314.

[32] K. J. Fidkowski, T. A. Oliver, J. Lu, and D. L. Darmofal, *p-Multigrid solution of high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations*, J. Comput. Phys., 207 (2005), pp. 92–113.

[33] E. H. Georgoulis and E. Süli, *Optimal error estimates for the hp-version interior penalty discontinuous Galerkin finite element method*, IMA J. Numer. Anal., 25 (2005), pp. 205–220.

[34] J. Gopalakrishnan and G. Kanschat, *A multilevel discontinuous Galerkin method*, Numer. Math., 95 (2003), pp. 527–550.

[35] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer Ser. Comput. Math. 4, Springer, Berlin, 1985.

[36] B. Heinrich and K. Pietsch, *Nitsche type mortaring for some elliptic problem with corner singularities*, Computing, 68 (2002), pp. 217–238.

[37] P. W. Hemker, W. Hoffmann, and M. H. van Raalte, *Two-level Fourier analysis of a multigrid approach for discontinuous Galerkin discretization*, SIAM J. Sci. Comput., 25 (2003), pp. 1018–1041.

[38] P. W. Hemker, W. Hoffmann, and M. H. van Raalte, *Fourier two-level analysis for discontinuous Galerkin discretization with linear elements*, Numer. Linear Algebra Appl., 11 (2004), pp. 473–491.

[39] P. Houston, C. Schwab, and E. Süli, *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, SIAM J. Numer. Anal., 39 (2002), pp. 2133–2163.

[40] C. Lasser and A. Toselli, *An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems*, Math. Comp., 72 (2003), pp. 1215–1238.

[41] H. Luo, J. D. Baum, and R. Löhner, *A p-multigrid discontinuous Galerkin method for the Euler equations on unstructured grids*, J. Comput. Phys., 211 (2006), pp. 767–783.

[42] B. S. Mascarenhas, B. T. Helenbrook, and H. L. Atkins, *Coupling p-multigrid to geometric multigrid for discontinuous Galerkin formulations of the convection-diffusion equation*, J. Comput. Phys., 229 (2010), pp. 3664–3674.

[43] C. R. Nastase and D. J. Mavriplis, *High-order discontinuous Galerkin methods using an hp-multigrid approach*, J. Comput. Phys., 213 (2006), pp. 330–357.

[44] L. N. Olson and J. B. Schroder, *Smoothed aggregation multigrid solvers for high-order discontinuous Galerkin methods for elliptic problems*, J. Comput. Phys., 230 (2011), pp. 6959–6976.

[45] I. Perugia and D. Schötzau, *An hp-analysis of the local discontinuous Galerkin method for diffusion problems*, J. Sci. Comput., 17 (2002), pp. 561–571.

[46] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Ser. Comput. Math. 23, Springer-Verlag, Berlin, 1994.

[47] J. Schöberl and C. Lehrenfeld, *Domain decomposition preconditioning for high order hybrid discontinuous Galerkin methods on tetrahedral meshes*, in Advance Finite Element Methods and Applications, Lect. Notes Appl. Comput. Mech. 66, Springer, New York, 2013, pp. 27–56.

[48] D. Schötzau, C. Schwab, and A. Toselli, *Mixed hp-DGFEM for incompressible flows*, SIAM J. Numer. Anal., 40 (2002), pp. 2171–2194.

[49] K. Shahbazi, D. J. Mavriplis, and N. K. Burgess, *Multigrid algorithms for high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations*, J. Comput. Phys., 228 (2009), pp. 7917–7940.

[50] B. Stamm and T. P. Wihler, *hp-optimal discontinuous Galerkin methods for linear elliptic problems*, Math. Comp., 79 (2010), pp. 2117–2133.

[51] M. H. van Raalte and P. W. Hemker, *Two-level multigrid analysis for the convection-diffusion equation discretized by a discontinuous Galerkin method*, Numer. Linear Algebra Appl., 12 (2005), pp. 563–584.