SIAM J. SCI. COMPUT.
Vol. 31, No. 5, pp. 3503–3528

SMOOTHED AGGREGATION MULTIGRID FOR THE DISCONTINUOUS GALERKIN METHOD*

F. PRILL†, M. LUKÁČOVÁ-MEDVIĎOVÁ‡, AND R. HARTMANN§

Abstract. The aim of this paper is to investigate theoretically as well as experimentally an algebraic multilevel algorithm for the solution of the linear systems that arise from the discontinuous Galerkin method. The smoothed aggregation multigrid, introduced by Vaněk for the conforming finite element method, is applied to low-order discretizations of convection-diffusion equations. For the elliptic model problem the algorithm is shown to be quasi-optimal. Adjustments for the case of nonvanishing advection, such as directionally implicit smoothing and a suitable splitting of the operator, are discussed. Several numerical experiments are presented for two-dimensional problems, including a Newton-type linearization of the compressible Navier–Stokes equations.

Key words. discontinuous Galerkin method, multigrid method, advection-diffusion equation, Navier–Stokes equations

AMS subject classifications. 35J05, 35J25, 35Q30, 65F10, 65M22, 65M55, 65N30, 65N35

DOI. 10.1137/080728457

1. Introduction.

The discontinuous Galerkin method (dGFEM) has become popular for the discretization of elliptic and hyperbolic partial differential equations for about a decade. To the desirable properties of this finite element method belong the flexibility in handling unstructured triangulations of complex geometries, mesh adaptation, and freedom in the choice of the polynomial basis. In addition, the dGFEM can be viewed as a generalization of the finite volume method, and recent studies have shown its suitability for the simulation of compressible flows [6, 21]. Moreover, the discretization possesses the variational background of the finite element method and is amenable to error analysis [3].

A major drawback is the fact that the linear algebraic systems arising from the dGFEM discretization consist of a relatively large number of unknowns and exhibit the customary numerical difficulties such as an increasing stiffness on large, anisotropic meshes. This motivates the development of efficient and scalable solvers, mostly belonging to the field of domain decomposition methods [1, 14, 28] or related defect-correction approaches, for example, the so-called p-multilevel algorithms [15]. These approaches reduce the systems to low-order subproblems that are usually discretized with piecewise constant or linear polynomials. It is the point of this paper to propose an optimal and scalable iterative scheme for these remaining algebraic systems arising from dGFEM discretizations of low order. We recall that the convergence rate of a

*Received by the editors June 25, 2008; accepted for publication (in revised form) June 11, 2009; published electronically September 4, 2009. The work of the first and third authors was partially supported by both the President's Initiative and Networking Fund of the Helmholtz Association of German Research Centres and the European project ADIGMA.
http://www.siam.org/journals/sisc/31-5/72845.html
†Corresponding author. Institute of Aerodynamics and Flow Technology, German Aerospace Center (DLR), D-38108 Braunschweig, Germany (florian.prill@dlr.de).
‡Institute of Numerical Simulation, Hamburg University of Technology, D-21071 Hamburg, Germany (lukacova@tu-harburg.de). The work of this author was partially supported by the Deutsche Forschungsgemeinschaft under grant LU 1470/1-1 and by the European Graduate School Differential Equations with Applications in Science and Engineering (DEASE), MEST-CT-2005-021122.
§Institute of Aerodynamics and Flow Technology, German Aerospace Center (DLR), D-38108 Braunschweig, Germany (ralf.hartmann@dlr.de).

(quasi-)optimal algorithm increases at most logarithmically with the size of a problem. The method is scalable if also the computational work for a problem containing n unknowns grows like $\mathcal{O}(n \log n)$.

Multigrid methods are known to be scalable for elliptic problems on regular grids. On structured meshes the grid hierarchy defined by successive global refinement steps can conveniently be employed for a multilevel method. The situation is different for unstructured grids and/or irregular domains. Here a mesh hierarchy is not provided a priori, and an alternative coarse level construction must be developed.

One possibility is to utilize a (not necessarily nested) family of triangulations of increasing grid size that has been generated automatically from the given geometry. However, especially in three dimensions automatic mesh generation is not a simple task. The second possibility is to construct the variational problems on grids that are defined purely algebraically. These approaches are termed *algebraic multigrid (AMG) methods* though besides the classical AMG method introduced in the early 1980s there exist a variety of other approaches exploiting the matrix data, such as element-based AMG (AMGe) [24], the algebraic multilevel iteration (AMLI) method [4], or (smoothed) aggregation multigrid [38, 39], the latter being the focus of this work.

The basic idea of smoothed aggregation multigrid is fairly intuitive and related to the geometric definition of coarser meshes. It consists in the construction of a nonoverlapping partition of the domain of interest, where each of the subsets constitutes a node on the next coarser level. The characteristic functions associated with this partition, however, are not suitable for the definition of a coarse basis. Therefore, their properties are improved by a smoothing procedure. Aggregation multigrid is closely related to the agglomeration multigrid algorithm which is in widespread use for the finite volume method [29]. This makes the approach especially attractive for the dGFEMs applied to convection-diffusion systems, where the hyperbolic terms are discretized by numerical fluxes as well.

For multigrid by smoothed aggregation restricted to elliptic model problems, there exists a convergence theory proving the optimality of the algorithm in terms of the mesh size with bounds on the condition number that are polynomial in the number of levels. It is based on the variational multigrid theory of Bramble [7] and requires only weak regularity assumptions, in contrast to classical convergence proofs for geometric multigrid [19]. The theory is closely related to the abstract convergence proofs for additive or multiplicative Schwarz methods [36].

In general, literature of the application of multilevel methods in the context of dGFEM is still not comprehensive. While being widely used and well analyzed for the conforming finite element method (see, for example, [10] and the references cited therein), multigrid for the dGFEM has been described so far by only a few researchers (see [9, 11, 16] for the geometric variant and [26] for a formulation of the AMLI method). Additionally, a less rigorous study using local Fourier analysis has been performed and published by Hemker, Hoffmann, and Van Raalte [22], apart from some results published for engineering applications. Another research focus has been the analysis of multilevel Schwarz methods. Similar to multigrid development, most coarse problems either have been based on meshes that are provided a priori or have been constructed as low-order discretizations while keeping the triangulation unchanged [14, 28]. An exception is the report by Lasser and Toselli [27], where the results for a two-level method utilizing an algebraic coarse problem have been presented.

In this paper we describe and analyze theoretically a smoothed aggregation multigrid for the dGFEM. This algorithm can be applied to elliptic problems on unstruc-

tured meshes. Additionally, we enrich it by components from geometric and finite volume multigrid, such as a line-implicit smoothing iteration. In particular, we use a Petrov–Galerkin variant that has been proposed in [17, 18]. This method is well suited for combination with standard agglomeration multigrid and less expensive in terms of memory requirements. The scalability is demonstrated by numerous experiments on structured and unstructured quadrilateral meshes. A comparison with standard geometric multigrid is performed as well where this is possible.

In addition to the model case of isotropic diffusion, we also apply the algorithm to nonsymmetric problems using a suitable splitting of the problem operator. We consider linear convection-diffusion and test the feasibility of the method as a preconditioner for the Newton-type implicit linearizations of nonlinear convection-diffusion. A typical benchmark are the two-dimensional compressible Navier–Stokes equations modelling the flow around the NACA 0012 airfoil.

The paper is structured as follows: In section 2 we briefly state the discontinuous Galerkin discretization of a scalar convection-diffusion model problem, along with some notation and the definition of the finite element spaces. Then, in section 3, we formulate the smoothed aggregation multigrid of Vaněk and coworkers and discuss some adjustments to the dGFEM model problem such as line-implicit smoothing and a separate treatment of the convective and diffusive terms. Section 4 is devoted to the convergence results of multigrid theory applied to the dGFEM. We present some numerical experiments in section 5 demonstrating the feasibility of the approach for purely diffusive and convection-diffusion problems. Finally, we draw conclusions in section 6.

2. Model problem and discretization. In the following we state the model problem and its discretization. We consider stabilization by interior penalties and by the method of Bassi and Rebay (see [6]).

2.1. Model problem. Let us consider the linear convection-diffusion equation on a bounded open polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with its boundary Γ being the union of its $(d-1)$-dimensional faces:

$$\mathcal{L}\,u \equiv -\nabla \cdot \big(\underline{a}\nabla u\big) + \beta \cdot \nabla u = f \quad \text{in } \Omega. \tag{2.1}$$

We assume that $f \in L^2(\Omega)$, $\underline{a} \in \mathbb{R}^{d\times d}$ is a symmetric positive-definite matrix, and $\beta = \{\beta_i\}_{i=1}^d$ is a vector field in $W^{1,\infty}(\Omega)$ such that $\nabla \cdot \beta \le 0$ in Ω. The existence of a unique solution for problem (2.1) equipped with homogeneous boundary conditions is well known; see, e.g., [34].

The boundary Γ is divided into sets $\Gamma = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}}$, such that the Dirichlet and Neumann boundary conditions can be imposed as

$$u = g_{\mathrm{D}} \text{ on } \Gamma_{\mathrm{D}}, \quad \mathbf{n} \cdot (\underline{a}\nabla u) = g_{\mathrm{N}} \quad \text{on } \Gamma_{\mathrm{N}},$$

assuming $\beta \cdot \mathbf{n} \ge 0$ on Γ_{N}. Here, by $\mathbf{n}(\mathbf{x})$, $\mathbf{x} \in \Gamma$, we denote the unit outward normal vector.

2.2. Meshes, trace operators, and finite element spaces. Let us assume that Ω can be subdivided into shape-regular meshes $\mathcal{T}_h$, consisting of convex open subsets (elements) $\kappa_j \neq \emptyset$, $j = 1, \dots, n_t$, of characteristic size $h > 0$:

$$\bar{\Omega} = \bigcup_{\kappa_j \in \mathcal{T}_h} \bar{\kappa}_j, \quad \kappa_i \cap \kappa_j = \emptyset, \quad 1 \le i, j \le n_t,\ i \neq j, \quad n_t := \operatorname{card} \mathcal{T}_h.$$

Throughout this paper we confine ourselves to partitionings into quadrilateral elements. In the theoretical results in section 4 we assume quasi uniformity; that is, the diameter h_κ of an element $\kappa \in \mathcal{T}_h$ is proportional to h. We define the set $\mathcal{E}$ of all interior and boundary faces, i.e., the smallest $(d-1)$-dimensional intersection between neighboring elements of the partition and between elements and the boundary Γ. Furthermore, we define the set of interior faces $\mathcal{E}_{\text{int}} = \{e \in \mathcal{E} \colon e \subset \Omega\}$ and its union $\Gamma_{\text{int}} := \{\mathbf{x} \in \Omega \colon \exists e \in \mathcal{E}_{\text{int}} \text{ with } \mathbf{x} \in e\}$. The dGFEM admits triangulations containing irregular vertices, also called hanging nodes, in a natural manner.

We define parametric finite elements based on the tesselation $\mathcal{T}_h$ in the usual way (see, for example, Quarteroni and Valli [34] for details), however, without continuity constraints on the interelement boundaries. The global discrete function space is given by

$$V_h^p := \left\{ v \in L^2(\Omega) \colon v\big|_\kappa \circ \sigma_\kappa \in \mathcal{Q}_p(\hat{\kappa}) \ \forall \kappa \in \mathcal{T}_h \right\},$$

where $\sigma_\kappa \colon \hat{\kappa} \to \kappa$ denotes a sufficiently smooth mapping from a reference element $\hat{\kappa} := [-1,1]^d$ to the triangulation and $\mathcal{Q}_p(\hat{\kappa})$ is the set of polynomials of degree less than or equal to p in each variable. We limit the focus of this paper to approximations with uniform polynomial degree in each element. As we are dealing with polynomials of low degree, piecewise (bi-)linear or constant, the specific choice of a basis for $\mathcal{Q}_p(\hat{\kappa})$, $n_\kappa := \dim(\mathcal{Q}_p(\hat{\kappa})) = (p+1)^d$, is less important now.

By $H^s(\mathcal{T}_h)$, $s \in \mathbb{R}^+$, we denote the broken Sobolev space, i.e., the space of functions on $\mathcal{T}_h$ whose restriction to an element $\kappa \in \mathcal{T}_h$ belongs to $H^s(\kappa)$. The associated broken norms and seminorms are defined as $\|u\|_{s,\mathcal{T}_h}^2 = \sum_{\kappa \in \mathcal{T}_h} \|u\|_{s,\kappa}^2$, $|u|_{s,\mathcal{T}_h}^2 = \sum_{\kappa \in \mathcal{T}_h} |u|_{s,\kappa}^2$.

For a function $v \in H^1(\mathcal{T}_h)$ we define v_κ^+, $\kappa \in \mathcal{T}_h$, to be the inner trace of v on $\partial\kappa$. The traces of functions in $H^1(\mathcal{T}_h)$ are double-valued for $\mathbf{x} \in \Gamma_{\text{int}}$, while on the boundary, $\mathbf{x} \in \Gamma$, the value $v(\mathbf{x})$ is unambiguous. For $\kappa \in \mathcal{T}_h$ with $\partial\kappa \setminus \Gamma \neq \emptyset$ there exists a neighboring element $\kappa' \in \mathcal{T}_h$ that shares a common edge $e = \bar{\kappa} \cap \bar{\kappa}' \in \mathcal{E}_{\text{int}}$ with κ. The outer trace v_κ^- of v on e is defined as the inner trace v_κ^+ relative to the element κ'.

Furthermore, let us define the following jump and average operators. For $v \in H^1(\mathcal{T}_h)$ the jump of v on the edge $e = \kappa \cap \kappa' \in \mathcal{E}_{\text{int}}$ is given by

$$[\![v]\!]_e \colon \left(\prod\nolimits_{\kappa \in \mathcal{T}_h} L^2(\partial\kappa) \right) \to \left[L^2(\Gamma \cup \Gamma_{\text{int}}) \right]^d, \quad [\![v]\!]_e = v_\kappa^+ \mathbf{n}_\kappa^+ + v_\kappa^- \mathbf{n}_\kappa^-,$$

where $\mathbf{n}_\kappa^+$ and $\mathbf{n}_\kappa^-$ denote the unit outward normal vectors to κ and κ', respectively. Additionally, the mean value of v on the edge e is defined as

$$\{v\}_e \colon \left(\prod\nolimits_{\kappa \in \mathcal{T}_h} L^2(\partial\kappa) \right) \to L^2(\Gamma \cup \Gamma_{\text{int}}), \quad \{v\}_e = \frac{1}{2} \left(v_\kappa^+ + v_\kappa^- \right).$$

For element boundaries $e \in \mathcal{E} \setminus \mathcal{E}_{\text{int}}$ that are part of the global domain boundary Γ the boundary values are defined unambiguously. We set

$$[\![v]\!]_e = v^+ \mathbf{n}, \quad \{v\}_e = v^+ \text{ on } e.$$

Finally, we introduce a jump notation for the inflow part $\partial_-\kappa$ of an element boundary that is defined as $\partial_-\kappa = \{\mathbf{x} \in \partial\kappa \colon \beta(\mathbf{x}) \cdot \mathbf{n} < 0\}$. Then the jump of u across $\partial_-\kappa \setminus \Gamma$ is defined by $[v]_\kappa := v_\kappa^+ - v_\kappa^-$. For the sake of simplicity the subscripts κ and e are omitted in the following.

2.3. The discontinuous Galerkin discretization. Arnold et al. [3] presented a unified formulation for the discontinuous Galerkin discretization of second-order operators which allows a simultaneous treatment of both the interior penalty stabilization and the approach of Bassi and Rebay (BR2 method).

The discrete problem is given in the following way: Find $u_h \in V_h^p$ such that

$$B_1(u_h, v_h) + B_2(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h^p \tag{2.2}$$

with bilinear forms B_1, B_2 for the diffusion and convection parts of problem (2.1), respectively. The latter uses a standard upwind term and is given by

$$\begin{aligned} B_2(u,v) := &\sum_{\kappa\in\mathcal{T}_h} \int_\kappa (\beta\cdot\nabla u)\, v \, d\mathbf{x} - \sum_{\kappa\in\mathcal{T}_h} \int_{\partial_-\kappa\setminus\Gamma} (\beta\cdot\mathbf{n})[u]v^+ \, ds \\ &- \sum_{\kappa\in\mathcal{T}_h} \int_{\partial_-\kappa\cap\Gamma} (\beta\cdot\mathbf{n})u^+v^+ \, ds, \\ \ell(v) := &- \sum_{\kappa\in\mathcal{T}_h} \int_{\partial_-\kappa\cap\Gamma_{\mathrm{D}}} (\beta\cdot\mathbf{n})g_{\mathrm{D}}v^+ \, ds + \sum_{\kappa\in\mathcal{T}_h} \int_\kappa fv \, d\mathbf{x}. \end{aligned}$$

The primal formulation of the elliptic operator is the following:

$$\begin{aligned} B_1(u,v) := &\int_\Omega \underline{a}\nabla_h u\cdot\nabla_h v \, d\mathbf{x} + \int_{\Gamma\cup\Gamma_{\mathrm{int}}} (\llbracket \hat{u}-u\rrbracket\cdot\{\underline{a}\nabla_h v\} - \{\hat{\mathbf{q}}\}\cdot\llbracket v\rrbracket) \, ds \\ &+ \int_{\Gamma_{\mathrm{int}}} (\{\hat{u}-u\}\llbracket\underline{a}\nabla_h v\rrbracket - \llbracket\hat{\mathbf{q}}\rrbracket\{v\}) \, ds, \end{aligned}$$

where we used the notation ∇_h for the broken gradient $(\nabla_h u)\big|_\kappa = \nabla\left(u\big|_\kappa\right)$, $\kappa\in\mathcal{T}_h$, and vector and scalar numerical fluxes on the boundary of κ,

$$\hat{\mathbf{q}}(u,\nabla u) \approx \underline{a}\nabla u, \qquad \hat{u}(u)\approx u.$$

Various choices of $\hat{\mathbf{q}}$, $\hat{u}$ give rise to a collection of consistent and stable dGFEMs [3]. Among the most popular schemes is the interior penalty method in its symmetric (SIPG) and nonsymmetric (NIPG) form, where we have

$$\begin{aligned} &\hat{u}_{\mathrm{SIPG}}(u) := \{u\}, && \hat{\mathbf{q}}_{\mathrm{SIPG}}(u,\nabla u) := \{\underline{a}\nabla u\} - \sigma\llbracket u\rrbracket, \quad \text{and} \\ &\hat{u}_{\mathrm{NIPG}}(u) := \{u\} + \mathbf{n}\cdot\llbracket u\rrbracket, && \hat{\mathbf{q}}_{\mathrm{NIPG}}(u,\nabla u) := \hat{\mathbf{q}}_{\mathrm{SIPG}}(u,\nabla u) \quad \text{on } \Gamma_{\mathrm{int}}, \end{aligned}$$

with appropriate modification at the boundary. The term $\sigma = \sigma(|\underline{a}|,h,p)$ denotes a stabilizing penalty parameter and is chosen as $\sigma\big|_e = \delta p^2 h_e^{-1}\|\sqrt{\underline{a}}\mathbf{n}_e\|^2_{L^\infty(e)}$ with a given constant $\delta>0$; see, e.g., [33].

The (modified) scheme of Bassi and Rebay (see [6]) uses the numerical fluxes

$$\hat{u}_{\mathrm{BR2}}(u) := \hat{u}_{\mathrm{SIPG}}(u), \quad \hat{\mathbf{q}}_{\mathrm{BR2}}(u,\nabla u) := \{\underline{a}\nabla u\} + \eta\{r_e(\llbracket u\rrbracket)\}, \tag{2.3}$$

where η is again a stabilizing constant and $r_e\colon \left[L^1(e)\right]^d \to \left[H^1(\mathcal{T}_h)\right]^d$ denotes a local lifting operator which is defined in a weak manner by

$$\int_\Omega r_e(\varphi)\cdot\tau \, d\mathbf{x} = -\int_e \varphi\cdot\{\underline{a}^T\tau\} \, ds \quad \forall\tau\in\left[H^1(\mathcal{T}_h)\right]^d, \ \varphi\in\left[L^1(e)\right]^d.$$

Using a local polynomial basis for $\mathcal{Q}_p(\hat{\kappa})$ together with the discretization (2.2), the discrete operator of the boundary value problem (2.1) is transformed into a stiffness matrix $\underline{A} \in \mathbb{R}^{n\times n}$, $n := \dim V_h^p = n_t n_\kappa$, while the functional $\ell(\cdot)$ takes the form of a right-hand side vector $\mathbf{b} \in \mathbb{R}^n$. Thus we are interested in solving a system of linear algebraic equations,

$$\underline{A}\mathbf{x} = \mathbf{b}.$$

Later, for the self-adjoint and coercive linear elliptic model problem, we will make use of the induced energy norm $\|\mathbf{u}\|_{\underline{A}} = (\mathbf{u},\mathbf{u})_{\underline{A}}$, where $(\mathbf{u},\mathbf{v})_{\underline{A}} = \mathbf{u}^T\underline{A}\mathbf{v}$, $\mathbf{u},\mathbf{v} \in \mathbb{R}^n$.

3. Description of the algorithm. In the following we state the classical multigrid algorithm and describe the smoothed aggregation approach for unstructured tesselations. Finally, we discuss the agglomeration heuristic and the smoothing iteration.

3.1. Basic algorithm. The smoothed aggregation multigrid can be viewed as a standard variational multigrid. We state the abstract algorithm [7] as a linear iteration $\mathbf{x} \leftarrow MG(\mathbf{x},\mathbf{b})$. This is equivalent to a preconditioner formulation,

$$\mathbf{x} \leftarrow \mathbf{x} - \underline{B}_{MG}\left(\underline{A}\mathbf{x}_n - \mathbf{b}\right), \quad \underline{B}_{MG}\mathbf{g} := MG(\mathbf{0},\mathbf{g}). \tag{3.1}$$

For the sake of simplicity, we focus on the V-cycle algorithm, although extensions as the W-cycle or full multigrid can be useful and are straightforward in implementation. Further we assume the model problem to be of purely elliptic type. The case of a nonvanishing advection term will be discussed in section 3.5.

We assume a sequence of finite-dimensional coarse spaces and operators

$$M_1 \subset M_2 \subset \cdots \subset M_J \equiv \mathbb{R}^n, \quad \left\{\underline{A}_j\right\}_{j=1}^J,\ \underline{A}_J \equiv \underline{A},$$

where M_1 denotes the coarsest problem space. Further we have prolongation operators $I_{k-1}^k \colon M_{k-1} \to M_k$ as well as restrictions $I_k^{k-1} \colon M_k \to M_{k-1}$, and smoothing iterations of fixed-point type

$$\mathbf{x} \leftarrow (\underline{I} - \underline{R}_k\underline{A}_k)\mathbf{x} + \underline{R}_k\mathbf{b} \tag{3.2}$$

with preconditioners $\underline{R}_k \colon M_k \to M_k$, $k = 2,\ldots,J$. For $i = 1,2,\ldots$, set $\underline{R}_k^{(i)} = \underline{R}_k$ if i is odd, and $\underline{R}_k^{(i)} = \underline{R}_k^T$ if i is even.

Inductively, we define $MG(\mathbf{x},\mathbf{b}) \equiv MG_J(\mathbf{x}^J,\mathbf{b}^J)$ as follows.

ALGORITHM 1: Multigrid algorithm $\mathbf{x}^l \leftarrow MG_l(\mathbf{x}^l,\mathbf{b}^l)$.

$MG_1(\mathbf{x}^1,\mathbf{b}^1) := \underline{A}_1^{-1}\mathbf{b}^1$ (*direct solution*)

for $l > 1$: Let $\mathbf{x}^l$, $\mathbf{b}^l \in M_l$ be given.

begin

Compute for $i = 1,\ldots,m$:

$$\mathbf{x}^l \leftarrow (\underline{I} - \underline{R}_l^{(i+m)}\underline{A}_l)\mathbf{x}^l + \underline{R}_l^{(i+m)}\mathbf{b}^l \ (\textit{presmoothing})$$

Set $\mathbf{b}^{l-1} = I_l^{l-1}\left(\mathbf{b}^l - \underline{A}_l\mathbf{x}^l\right)$.

$\mathbf{x}^l \leftarrow \mathbf{x}^l + I_{l-1}^l\mathbf{x}^{l-1}$, where $\mathbf{x}^{l-1} := MG_{l-1}(\mathbf{0},\mathbf{b}^{l-1})$ (*coarse-grid correction*)

Compute for $i = m+1,\ldots,2m$:

$$\mathbf{x}^l \leftarrow (\underline{I} - \underline{R}_l^{(i+m)}\underline{A}_l)\mathbf{x}^l + \underline{R}_l^{(i+m)}\mathbf{b}^l \ (\textit{postsmoothing})$$

end

The specific choice of the subspaces M_k, $k = 1,\ldots,J$, as well as the level operators $\underline{A}_k$ and the transfer operators will be described next. The construction of the smoothing iteration $\underline{R}_k$ is discussed in section 3.4.

3.2. Smoothed aggregation coarse spaces. Now we describe the smoothed aggregation multigrid [38, 39]. Our starting point is a nonoverlapping partition of the domain into strongly connected subsets a_i^k, $k = 1, \dots, J$ fixed, $i = 1, \dots, n_k$, termed *agglomerates*. Each agglomerate a_i^k on a given level k is itself formed by fusing together fine level elements in $\left\{a_j^{k+1}\right\}_{j=1}^{n_{k+1}}$, where we formally identify the elements $\kappa \in \mathcal{T}_h$ with the agglomerates on the finest level.

Equivalent to the element agglomeration process, a partitioning of the index sets $\mathcal{I}^k$, $k = 1, \dots, J$, of the degrees of freedom takes place. Groups of fine level degrees of freedom $\mathcal{I}_l^k \subseteq \mathcal{I}^k$, $l = 1, \dots, n_k$, are *aggregated* and give rise to a common degree of freedom on the coarse level $k-1$. This is realized recursively, where the particular agglomeration strategy will be described in section 3.3 below.

The following notation will be of later use. We define the composite set $\mathcal{I}_l^{k,J}$, $k = 1, \dots, J$, $l = 1, \dots, n_k$, of fine level indices by

$$\mathcal{I}_l^{k,J} := \left\{ i_J \in \mathcal{I}^J : \text{there exist indices } (i_k, \dots, i_J) \text{ with } i_s \in \mathcal{I}_{i_{s-1}}^s, k < s \le J, i_k = l \right\};$$

i.e., $\mathcal{I}_l^{k,J}$ denotes the set of fine level indices corresponding to basis functions with support in a_l^k.

The characteristic functions $\chi_{a_i^k}$, corresponding to the agglomerates on level k, span a linear space which can be used as the ansatz space for a Galerkin procedure forming the coarse solution $\tilde{u}_k$ for the problem

$$B_1(\tilde{u}_k, \tilde{v}_k) = f(\tilde{v}_k) \quad \forall \tilde{v}_k \in \operatorname{span}_{i \in \mathcal{I}^k} \chi_{a_i^k}.$$

In view of the matrix formulation this defines the so-called tentative prolongation and restriction

$$\tilde{I}_{k-1}^k \in \mathbb{R}^{n_k \times n_{k-1}}, \quad \left(\tilde{I}_{k-1}^k\right)_{ij} := \chi_{a_j^k}(\mathbf{x}_i), \quad \tilde{I}_k^{k-1} := \left(\tilde{I}_{k-1}^k\right)^T,$$

where $\mathbf{x}_i$ denotes the geometric location of the ith unknown in M_k. In fact, the coarse degrees of freedom are simply constructed as the sums of the fine level unknowns inside the element cluster. In order to construct the interpolation operators we restrict ourselves to orthonormalized zero-one prolongations, though more general choices are possible [38].

It can be shown by an argument on the matrix graph that the maximum number of nonzero entries per row in the coarse stiffness matrices

$$\underline{\tilde{A}}_k = \left(\tilde{I}_k^J\right)^T \underline{A} \tilde{I}_k^J, \text{ with } \tilde{I}_k^J := \tilde{I}_{J-1}^J \cdots \tilde{I}_{k+1}^{k+2} \tilde{I}_k^{k+1}, \quad k = 1, \dots, J-1, \tag{3.3}$$

is bounded under the condition that the partitioning algorithm keeps the number of direct neighbors of the agglomerates approximately constant [18]. This assures that the number of nonzeros decreases with the coarsening factor that is necessary for the scalability of the algorithm. The approximation properties of the tentative transfer operators, however, are less satisfactory. This fact has been observed experimentally [30], and it is supported by the regularity-free convergence theory.

To provide a stable coarse basis we apply symmetric smoothing polynomials $s^k : \mathbb{R}^{n_k \times n_k} \to \mathbb{R}^{n_k \times n_k}$, $1 \le k \le J$, to the piecewise constant transfer operators. Following Guillard and Vaněk [18] we employ different smoothers s_L^k, s_R^k for the restriction and the prolongation, respectively,

$$I_k^{k-1} := \left(s_L^k(\underline{A}_k) \tilde{I}_{k-1}^k\right)^T, \quad I_{k-1}^k := s_R^k(\underline{A}_k) \tilde{I}_{k-1}^k.$$

Then, defining the coarse level matrices as

$$\underline{A}_{k-1} = I_k^{k-1} \underline{A}_k I_{k-1}^k$$

is equivalent to the Galerkin procedure $\underline{A}_{k-1} = \left[B_1(\phi_i^{k-1}, \psi_j^{k-1})\right]_{i,j=1}^{n_{k-1}}$ with coarse basis functions ϕ_i^{k-1}, ψ_j^{k-1}, which are inductively defined as

$$\begin{bmatrix} \phi_1^{k-1} \\ \vdots \\ \phi_{n_{k-1}}^{k-1} \end{bmatrix} = \left(\tilde{I}_{k-1}^k\right)^T s_L(\underline{A}_k) \begin{bmatrix} \phi_1^k \\ \vdots \\ \phi_{n_k}^k \end{bmatrix}, \quad \begin{bmatrix} \psi_1^{k-1} \\ \vdots \\ \psi_{n_{k-1}}^{k-1} \end{bmatrix} = \left(\tilde{I}_{k-1}^k\right)^T s_R(\underline{A}_k) \begin{bmatrix} \psi_1^k \\ \vdots \\ \psi_{n_k}^k \end{bmatrix}.$$

The smoothing polynomials produce overlap between the supports of the coarse functions; this is especially noteworthy for the dGFEM, where the tentative coarse functions possess no overlap at all. However, the smoothed coarse functions remain piecewise discontinuous.

We choose between two variants of transfer operator smoothers.

1. The (symmetric) *Ritz–Galerkin* multigrid is given by

$$s_L^k(\underline{A}_k) = s_R^k(\underline{A}_k) := \underline{I} - \frac{4}{3\lambda_k}\underline{A}_k, \tag{3.4}$$

where $\lambda_k \geq \varrho(\underline{A}_k)$ is an approximation for the spectral radius.

2. The *Petrov–Galerkin* approach is defined by

$$s_L^k(\underline{A}_k) := \underline{I}, \quad s_R^k(\underline{A}_k) := \underline{I} - \frac{1}{\lambda_k}\underline{A}_k. \tag{3.5}$$

Other choices for optimal smoothing polynomials are given in [17]. The degree of the smoothing polynomials is closely related to the diameter (coarsening ratio) of the agglomerates a_i^k. While an increased degree improves the smoothing effect, the overlap increases, too. Therefore fill-in optimality of the level matrices demands either a smaller degree or larger agglomerates. Thus the smoothing polynomial must be carefully chosen to match the coarsening ratio. In the presence of additional convective terms yet another argument takes place: As hyperbolic problems may exhibit solutions with discontinuities, the preferable multigrid coarsening factor would be as small as possible to represent these local phenomena on coarser levels. Traditional finite-volume multigrid is based on a coarsening ratio of 2 to resolve hyperbolic features of the solution.

The tentative operators $\tilde{I}_{k-1}^k$ as well as the smoothed prolongations I_{k-1}^k and the level matrices $\underline{A}_k$ are computed in a setup phase. In addition, with a diagonal scaling, we achieve that $(\tilde{I}_{k-1}^k)^T(\tilde{I}_{k-1}^k) = \underline{I}$, which is necessary for the convergence theory presented by Theorem 4.5.

For completeness, we explicitly define the hierarchy of coarse spaces M_k, $k = 1, \ldots, J-1$, for the variational multigrid [17, 38]. The space M_k is given as

$$M_k := \text{Range}\left(I_k^J\right), \quad k = 1, \ldots, J-1,$$

where, for the Ritz–Galerkin variant, we define I_k^J analogously to (3.3) as

$$I_k^J = s_R^J(\underline{A}_J)\tilde{I}_{J-1}^J \cdots s_R^{k+1}(\underline{A}_{k+1})\tilde{I}_k^{k+1}, \quad k = 1, \ldots, J-1.$$

For the Petrov–Galerkin multigrid, we set

$$I_k^J = \left(s_R^J(\underline{A}_J)\right)^{\frac{1}{2}} \tilde{I}_{J-1}^J \cdots \left(s_R^{k+1}(\underline{A}_{k+1})\right)^{\frac{1}{2}} \tilde{I}_k^{k+1}, \quad k = 1, \ldots, J-1.$$

M_k is equipped with the scalar product $(\cdot,\cdot)_k$, given by $(\mathbf{u},\mathbf{v})_k = (\mathbf{x},\mathbf{y})_{R^{n_k}}$, where $\mathbf{u} = I_k^J\mathbf{x}$, $\mathbf{v} = I_k^J\mathbf{y}$, and $\mathbf{x},\mathbf{y}$ are vectors orthogonal to the nullspace $\mathrm{Ker}\left(I_k^J\right)$.

3.3. Agglomeration strategy. Partitionings of the domain into connected sets of neighboring elements are required by many hierarchical solution strategies, for example, domain decomposition methods, agglomeration finite volume multigrid, and algebraic multilevel approaches like AMGe. Thus, there exists a variety of coarsening algorithms in the literature [18, 23, 24, 25, 30, 31, 38].

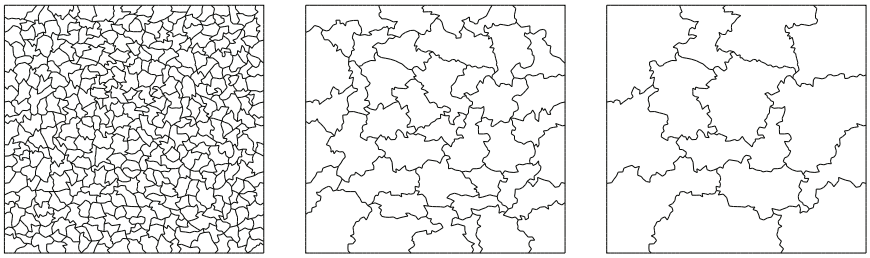
Many approaches heuristically approximate a maximum independent set in the corresponding node graph and are of ideally linear complexity. In the current work we use a greedy clustering algorithm that operates on the face-face graph of the tesselation, i.e., the undirected graph whose vertices are faces of the triangulation with edges between faces sharing a common triangulation vertex. The algorithm was initially described by Jones and Vassilevski [24] in the context of the AMGe method.

The basic idea behind the two-dimensional algorithm is to identify each face of the triangulation with a weighting scalar that is incremented during the iterations whenever neighboring faces are processed. The algorithm then tries to build monotonically increasing sequences of face weights, while adding the adjacent elements to the new agglomerate. More precisely, for a given level index m the algorithm can be formulated as follows:

Initialize a weighting vector w, $w(e) \leftarrow 0$ for all faces $e \in \mathcal{E}^m$; set $n \leftarrow 0$.
while *faces left* **do** global search
  Find a face $e_k \in \mathcal{E}^m$ with maximum weight $w(e_k) \geq 0$.
  Set $n \leftarrow n+1$, $a_n^m \leftarrow \emptyset$.
  while *complete* = *false* **do** construction of aggregate a_n^m
    Set $a_n^m \leftarrow a_n^m \cup (a_p^{m+1} \cup a_q^{m+1})$, where $e_k = a_p^{m+1} \cap a_q^{m+1}$,
      and set $w_{\max} \leftarrow w(e_k)$, $w(e_k) \leftarrow -1$.
    Set $w(e_i) \leftarrow w(e_i) + 1$ for all faces e_i with $w(e_i) \neq -1$,
      where e_i is a neighbor of e_k.
    Set $w(e_j) \leftarrow w(e_j) + 1$ for all faces e_j, with $w(e_j) \neq -1$, where
      e_j is a neighbor of e_k, and e_j, e_k are faces of a common element.
    From the neighbors of e_k choose a face $e_{k'}$ with maximum weight $w(e_{k'})$
      that shares a common element with e_k.
    if $w(e_k) < w_{\max}$ **then**
      Set $w(e) \leftarrow -1$ for all faces e of elements in a_n^m.
    else
      $k \leftarrow k'$; *complete* $\leftarrow$ *true*.

Compared to [24] the choice of the face $e_{k'}$ with maximum weight is slightly modified to achieve strongly connected agglomerates; see [25]. Note that the algorithm requires the definition of a topology on coarser levels, containing elements a_i^m and faces $\mathcal{E}^m$, $m < J$. The method can be extended to three-dimensional problems by operating on the side-side graph [25].

The algorithm has a typical coarsening ratio of 2 in each spatial direction, and it is able to rebuild the natural coarse hierarchy of $J = 1 - \log_2 h$ levels in globally refined structured meshes. On unstructured tesselations, however, the shape of the element clusters may deteriorate; cf. Figure 3.1. Furthermore, from the point of view of optimal complexity, the agglomerates constructed by the algorithm are suitable for



(a) Level J-1 (b) Level J-3 (c) Level J-4

FIG. 3.1. *Example of an agglomeration on an unstructured mesh with* 5316 *elements. The lines denote the boundaries of the agglomerates* a_i^k. *See Figure* 5.1(a) *for a plot of the underlying tesselation.*

the Petrov–Galerkin multigrid but not well matched for the Ritz–Galerkin multigrid. As explained in section 3.2, the ratio of the number of fine and coarse level degrees of freedom must increase with the degrees of the smoothing polynomials in order to preserve the scalability of the method. More precisely, the sum of the polynomial degrees of the transfer operator smoothers, which is $r = 2$ for the Ritz–Galerkin multigrid, requires a coarsening factor of 3 to keep the nearest neighbor stencil property. Throughout the numerical tests in section 5, however, a coarsening factor of 2 was employed, which is appropriate for the Petrov–Galerkin method.

3.4. Smoothing iteration. In this section we discuss the multilevel smoothing iteration (3.2) in more detail. The operators $\underline{R}_k$, $k = 2, \ldots, J$, were chosen to be of Jacobi or Gauß–Seidel type. In particular, a directionally implicit variant was implemented. From an abstract point of view the so-called line-implicit relaxation schemes belong to the class of additive or multiplicative iterative schemes associated with a subspace decomposition $M_k = \sum_{i=1}^{l} M_k^i$. We assume projections $\underline{Q}_k^i : M_k \to M_k^i$ with respect to the inner product $(\cdot, \cdot)_k$. The additive (Jacobi) smoother is then defined by

$$\underline{R}_k = \gamma \sum_{i=1}^{l} \underline{A}_{k,i}^{-1} \underline{Q}_k^i, \quad k = 2, \ldots, J,$$

with a damping factor $\gamma > 0$ and exact local problems $\underline{A}_{k,i} : M_k^i \to M_k^i$, $i = 1, \ldots, l$, given by $(\underline{A}_{k,i}\mathbf{v}, \mathbf{w})_k = (\underline{A}_k\mathbf{v}, \mathbf{w})_k$ for all $\mathbf{w} \in M_k^i$. The multiplicative (Gauß–Seidel) variant is the result of the following procedure:

for $i = 1, \ldots, l$ **do**
$\mathbf{v}_i = \mathbf{v}_{i-1} + \underline{A}_{k,i}^{-1} \underline{Q}_k^i (\mathbf{b} - \underline{A}_k \mathbf{v}_{i-1})$, where $\mathbf{v}_0 := \mathbf{0}$.
$\underline{R}_k \mathbf{b} = \mathbf{v}_l$.

For the case that the space of unknowns is decomposed into a direct sum, the above methods reduce to classical block Jacobi or Gauß–Seidel schemes applied to the stiffness matrix. Obviously the subspace decomposition determines the additive and multiplicative scheme.

The choice of the spaces is much less restricted for multilevel smoothers than for the related classical Schwarz methods. We set the following assumption.

ASSUMPTION 1 (smoothing property). *We assume that there exist constants* $C_R > 0$, $\theta \in [0, 2)$, *independent of the level* k, *such that the matrices* $\underline{R}_k$, $k =$

$2, \ldots, J$, *satisfy*

$$\frac{\|\mathbf{u}\|_2^2}{\varrho(\underline{A}_k)} \leq C_R^2 \left(2(\underline{R}_k \mathbf{u}, \mathbf{u}) - (\underline{A}_k \underline{R}_k \mathbf{u}, \underline{R}_k \mathbf{u})\right), \quad k = 2, \ldots, J,$$

$$\text{and} \qquad (\underline{R}_k \underline{A}_k \mathbf{u}, \underline{R}_k \underline{A}_k \mathbf{u})_{\underline{A}_k} \leq \theta \, (\underline{R}_k \underline{A}_k \mathbf{u}, \mathbf{u})_{\underline{A}_k} \quad \forall \mathbf{u} \in \mathbb{R}^{n_k}. \tag{3.6}$$

Remark 1. Assumption 1 must be modified in the following way for the Petrov–Galerkin multigrid [18, 23]: The algorithm can be reformulated as a Ritz–Galerkin approach with smoothing matrix $\tilde{s}^k(\underline{A}_k) := \left(s_R^k\right)^{\frac{1}{2}} (\underline{A}_k)$ with the help of an additional presmoothing step

$$\mathbf{x} \leftarrow \tilde{s}^k(\underline{A}_k)\mathbf{x} + (\underline{I} - \tilde{s}^k(\underline{A}_k))\underline{A}_k^+ \mathbf{b}, \tag{3.7}$$

where $\underline{A}_k^+$ denotes the pseudoinverse. Therefore, to apply the abstract convergence theory, (3.6) must hold for the modified smoother $\underline{R}_k'$:

$$\underline{K}_k' = \tilde{s}^k(\underline{A}_k)(\underline{I} - \underline{R}_k \underline{A}_k), \quad \underline{R}_k' = (\underline{I} - \underline{K}_k')\underline{A}_k^+.$$

The additive and the multiplicative smoother both satisfy Assumption 1 under weak conditions [7]. In particular, these are fulfilled when the triangulation is decomposed into sequences of elements (*lines*) without loops or cycles. This is a popular smoothing scheme since the local solvers $\underline{A}_{k,i}^{-1}$ are (block) tridiagonal linear systems that can be efficiently factorized by LU decomposition and solved via forward and backward substitution.

Whereas from the theory the particular way of line decomposition does not need to fulfill further properties, from a practical point of view it creates a difficult task. It is observed experimentally, and has been supported by local Fourier analysis [31], that lines should follow strong couplings in the stiffness matrix when the iterative scheme is used as a multigrid smoother. In the case of pure convection the smoother ideally becomes an exact solver. This is a common strategy for the construction of a robust multigrid method that is insensitive with respect to the strength $\|\beta\|_\infty$ of the advection term.

We make use of a symmetric coupling matrix $\underline{C} \in \mathbb{R}^{n_t \times n_t}$, where $n_t = \operatorname{card} \mathcal{T}_h$, $c_{ij} \in [0, 1]$, which measures interdependencies between the elements of the triangulation. Elements κ_i, $\kappa_j \in \mathcal{T}_h$ with $c_{ij} \approx 1$ are termed *strongly coupled*. We need to formulate an appropriate scalar model problem $\tilde{\mathcal{L}}\,\tilde{u} = \tilde{f}$, where $\tilde{\mathcal{L}}$ is a convection-diffusion operator analogous to (2.1). When dealing with a linear primal problem $\mathcal{L}\,u = f$, the same convection and diffusion coefficients can be chosen for $\tilde{\mathcal{L}}$. The task is less trivial for nonlinear convection-diffusion, e.g., the Navier–Stokes equations. There, one can extract the vector β from the nonlinear state. However, the diffusion tensor $\underline{a}$ must be chosen a priori.

Let us define the matrix $\underline{C}$ as the stiffness matrix $\underline{A}_{\tilde{\mathcal{L}}}$, normalized as follows:

$$\underline{C} = \{c_{ij}\}_{i,j=1}^{n_t}, \quad c_{ij} := \frac{\tilde{c}_{ij}}{\max_{1 \leq i \leq n_t} \tilde{c}_{ij}}, \quad 1 \leq i, j \leq n_t,$$
$$\text{with } \tilde{c}_{ij} = \max\left(\left|\left(\underline{A}_{\tilde{\mathcal{L}}}\right)_{ij}\right|, \left|\left(\underline{A}_{\tilde{\mathcal{L}}}\right)_{ji}\right|\right) \approx \max\left(\left|\frac{\partial \tilde{\mathcal{L}}_i}{\partial \phi_j}\right|, \left|\frac{\partial \tilde{\mathcal{L}}_j}{\partial \phi_i}\right|\right).$$

With an appropriate coupling matrix at hand we use a greedy line construction heuristic to extract the subspaces M_k^i, $i = 1, \ldots, l$, from the triangulation, where l denotes the number of lines. The algorithm is similar to the one described by Okusanya (see [15, 31]). Starting with a seed element, the neighboring element with maximum coupling measure is attached and forms a line with two elements. This happens successively until the line is complete, i.e., until there exist no unprocessed elements coupled to the end of the current line with a coupling measure above a given threshold (chosen as $c_{ij} \geq 0.95$). Before appending an element, the algorithm verifies that the line still results in a block-tridiagonal linear system, i.e., that the line does not contain loops. The whole procedure is performed twice for each seed element such that the resulting line stretches in both directions. A typical result is shown in Figure 5.1(b). The same strategy can be applied to the triangulation as well as to the agglomerated levels, where we define

$$\underline{C}^{k-1} = \tilde{I}_k^{k-1} \underline{C}^k \tilde{I}_{k-1}^k, \quad k = 2, \ldots, J, \quad \underline{C}^J := \underline{C}. \tag{3.8}$$

Remark 2. The described line construction procedure reproduces the anisotropic features of the equations only to a certain degree, since it employs a scalar approximation $\tilde{\mathcal{L}}$ and a heuristic depending on the element numbering. Moreover, the coupling matrices created by (3.8) do not completely reflect the character of the coarse matrices $\underline{A}_j$, because they are merely formed by the transfer operators of the convection part. A fully matrix-dependent coupling criterion can, for example, be found in [32].

3.5. Treatment of convective terms. The smoothed aggregation technique has been described and analyzed by Vaněk, Brezina, and Mandel [38] for elliptic problems. The discretization of the convection-diffusion problem (2.1), however, also includes a hyperbolic term whose upwind character deteriorates under the smoothed interpolation. Following Guillard and Vaněk [18] we therefore split the discrete operator

$$\underline{A}\mathbf{x} = (\underline{A}^v + \underline{A}^c)\,\mathbf{x}$$

into its diffusive and convective components $\underline{A}^v$ and $\underline{A}^c$, respectively. Then, the coarse level matrices for the convective term are constructed using the simple, nonsmoothed transfer operators $\tilde{I}_{k-1}^k$, $\tilde{I}_k^{k-1}$. It can be shown that aggregation with piecewise constant interpolation is a sufficient approximation of the rediscretized operator on the agglomerates; see, e.g., [18].

4. Convergence bounds. The convergence estimate for the smoothed aggregation multigrid requires some theoretical tools that are presented in section 4.1. Throughout this section we consider an elliptic model problem, i.e., $\underline{a} = \underline{I}$, $\beta \equiv \mathbf{0}$ in (2.1), discretized by the bilinear form $B_1(\cdot, \cdot)$. We further confine ourselves to the symmetric discretizations BR2 and SIPG and postpone until section 4.3 the discussion of the NIPG. By C, C_i we will denote generic constants.

4.1. Theoretical tools. The following inequalities for piecewise polynomial functions can, for example, be found in [2, 33, 34].

LEMMA 4.1. *There exist constants $C_1, C_2 > 0$ depending on the shape regularity constant of $\mathcal{T}_h$ and the polynomial degree p, such that for all $\kappa \in \mathcal{T}_h$ and $\phi \in V_h^p$ the following inequalities hold:*

$$|\phi|_{1,\kappa}^2 \leq C_1 h_e^{-2} \|\phi\|_{0,\kappa}^2 \qquad \textit{(local inverse inequality)}, \tag{4.1}$$

$$\|\phi\|_{0,e}^2 \leq C_2 h_e^{-1} \|\phi\|_{0,\kappa}^2, \quad e \subset \partial\kappa \qquad \textit{(local trace inequality)}. \tag{4.2}$$

LEMMA 4.2 (Poincaré–Friedrichs inequality [2]). *Let $D \subseteq \Omega$ be an open connected polyhedral domain that is the union of some elements in $\mathcal{T}_h$. Then there exists $C > 0$, depending on $\mathcal{T}_h$ and the shape of D, such that for all $u \in H^1(\mathcal{T}_h)$*

$$\|u - \bar{u}\|_{0,D}^2 \leq C \operatorname{diam}(D)^2 \left(\sum_{\kappa \in \mathcal{T}_h, \kappa \subset D} |u|_{1,\kappa}^2 + \sum_{e \in \mathcal{E}, e \subset D} h_e^{-1} \|[\![u]\!]\|_{0,e}^2 \right), \tag{4.3}$$

where $\bar{u} := \frac{1}{\mu(D)} \int_D u \, d\mathbf{x}$.

We also recall the following lemma stating coercivity and boundedness of the bilinear form $B_1(\cdot,\cdot)$; see, e.g., [3].

LEMMA 4.3. *With respect to the norm $\|\!|\cdot|\!\|_{dG} \colon V_h^p \to \mathbb{R}^+$, defined by*

$$\|\!|v|\!\|_{dG}^2 = \sum_{\kappa \in \mathcal{T}_h} |v|_{1,\kappa}^2 + \sum_{e \in \mathcal{E}} h_e^{-1} \|[\![v]\!]\|_{0,e}^2, \quad v \in V_h^p,$$

the bilinear form $B_1(\cdot,\cdot)$ is coercive; i.e., there exists $C > 0$ such that

$$B_1(u,u) \geq C \|\!|u|\!\|_{dG}^2 \quad \forall u \in V_h^p.$$

Further, there exists $C_B > 0$ such that

$$B_1(u,v) \leq C_B \|\!|u|\!\|_{dG} \|\!|v|\!\|_{dG} \quad \forall u, v \in V_h^p.$$

Remark 3. To ensure positive-definiteness in the case of the SIPG and the BR2 method, the amount of the discontinuity-penalization given by δ and η in section 2.2 must be chosen sufficiently large. While for the BR2 scheme one needs to assume $\eta \geq 2d$, lower bounds for the SIPG parameter δ are mesh-dependent [35].

Finally, we have a bound for the spectral radius of the stiffness matrix, similar to that of the standard finite element method.

LEMMA 4.4. *For the spectral radius $\varrho(\underline{A})$, where $\underline{A} \in \mathbb{R}^{n \times n}$ denotes the stiffness matrix corresponding to $B_1(\cdot,\cdot)$, we have*

$$\varrho(\underline{A}) \leq C h_{\min}^{d-2}, \quad h_{\min} := \min_{\kappa \in \mathcal{T}_h} h_\kappa, \tag{4.4}$$

with a constant $C > 0$ depending on the stabilization parameters δ, η, the shape-regularity of $\mathcal{T}_h$, and the polynomial degree p of V_h^p.

Proof. The proof relies on a bound for the energy norm $u \mapsto B_1(u,u)$. We have

$$B_1(u,u) = |u|_{1,\mathcal{T}_h}^2 - 2 \int_{\Gamma \cup \Gamma_{\text{int}}} [\![u]\!] \cdot \{\nabla_h u\} \, ds - s(u) \quad \text{for } u \in V_h^p, \tag{4.5}$$

with a stabilization term $s(u) := \sum_{e \in \mathcal{E}} \delta h_e^{-1} p^2 \int_e [\![u]\!]^2 ds$ for the interior penalty method and $s(u) := \sum_{e \in \mathcal{E}} \eta \int_\Omega \left(r_e([\![u]\!])\right)^2 ds$ for the BR2 method. Using the property of the lifting operator r_e,

$$C_1 h_e^{-1} \|u\|_{0,e}^2 \leq \|r_e(u)\|_{0,\Omega}^2 \leq C_2 h_e^{-1} \|u\|_{0,e}^2, \quad e \in \mathcal{E},$$

it suffices to consider the SIPG stabilization. Applying Young's inequality together with (4.2) we get

$$\int_{\Gamma \cup \Gamma_{\text{int}}} [\![u]\!] \cdot \{\nabla_h u\} \, ds \leq \frac{1}{2h_e} \int_{\Gamma \cup \Gamma_{\text{int}}} [\![u]\!]^2 \, ds + \frac{C}{2} |u|_{1,\mathcal{T}_h}^2. \tag{4.6}$$

Substituting (4.6) in (4.5) we have

$$B_1(u,u) \leq (1+C)|u|^2_{1,\mathcal{T}_h} + (1+\delta p^2)\int_{\Gamma\cup\Gamma_{\text{int}}} h_e^{-1}[\![u]\!]^2 \, ds.$$

Applying (4.1), (4.2) once again we get the bound

$$B_1(u,u) \leq Ch_{\min}^{-2}\|u\|^2_{0,\mathcal{T}_h},$$

with C depending on δ, p. The bound for the spectral radius of $\underline{A}$ then follows from the equivalence between the discrete norm $\|\cdot\|_2$ and $\|\cdot\|_{0,\mathcal{T}_h}$: Let $\mathbf{u} \in \mathbb{R}^n$ denote the vector of coefficients for $u = \sum_{i=1}^n u_i\phi_i$. Then we have

$$\varrho(\underline{A}) \leq \frac{\mathbf{u}^T\underline{A}\mathbf{u}}{\|\mathbf{u}\|_2^2} \leq \frac{B_1(u,u)}{\|\mathbf{u}\|_2^2} \leq Ch_{\min}^{-2}\frac{\|u\|^2_{0,\mathcal{T}_h}}{\|\mathbf{u}\|_2^2} \leq Ch_{\min}^{d-2},$$

which concludes the proof of the lemma. ∎

4.2. Convergence estimate. Our V-cycle estimate is an application of the following convergence theorem.

THEOREM 4.5 (Vaněk, Brezina, and Mandel [38], Guillard, Janka, and Vaněk [17]). *Assume a quasi-uniform family of meshes $\mathcal{T}_h$, together with orthonormal piecewise constant operators $\tilde{I}_{k-1}^k$, $\tilde{I}_k^J$ as defined in* (3.3). *Let the prolongation and restriction smoothers be given by the polynomials* (3.4) *and* (3.5) *for the Ritz–Galerkin and the Petrov–Galerkin multigrid, respectively.*

We assume $\lambda_l := (r+1)^{2(l-J)}\lambda$, $\lambda \geq \varrho(\underline{A})$, where $r = 2$ in the Ritz–Galerkin multigrid, and $r = 1$ in the Petrov–Galerkin case. Further let the smoothers $\underline{R}_k$, $k = 2,\ldots,J$, satisfy Assumption 1 *with constants $C_R > 1$, $\theta \in [0,2)$. Then*

$$\|\underline{A}^{-1}\mathbf{b} - MG(\mathbf{x},\mathbf{b})\|_{\underline{A}} \leq \left(1 - \frac{1}{C}\right)\|\underline{A}^{-1}\mathbf{b} - \mathbf{x}\|_{\underline{A}}, \tag{4.7}$$

where

$$C = \left[2 + C_1\frac{J-1}{r+1} + C_R\left(C_1 + C_2 + C_1C_2\frac{J-1}{r+1}\right)\sqrt{\frac{\theta}{2-\theta}}\right]^2\frac{J-1}{2-\theta} \tag{4.8}$$

with C_1 independent of the number of levels, and $C_2 := \frac{4}{3}$ in the Ritz–Galerkin case and $C_2 := \sqrt{2+\frac{\pi^2}{3}}$ for the Petrov–Galerkin multigrid. If the multigrid algorithm is used as a preconditioner, then it follows from the Rayleigh quotient characterization of the extreme eigenvalues that there holds $\operatorname{cond}(\underline{B}_{MG}\underline{A}) \leq C$.

The estimate is based on an abstract convergence result proved in [7, 8]. In the "nested-inherited regularity-free" approach of Bramble and Xu the multigrid algorithm is interpreted as a multiplicative subspace correction method. It relies on two conditions, first, a smoothing property formulated in Assumption 1 and, second, a weak approximation property. For the latter, by virtue of a lemma by Vaněk, Brezina, and Mandel [38], it is sufficient to verify assumptions on the prolongation smoother and the linear spaces constructed by the piecewise constant transfer operators. The properties of the smoothing polynomials chosen in (3.4), (3.5) are investigated in [17, 38], and their extremal values give rise to the constant C_2 in (4.8).

Regarding the disjoint element agglomerates and the corresponding piecewise constant transfer operators, it has to be verified that there exist linear mappings

$$\underline{\tilde{Q}}_l \colon \mathbb{R}^{n_J} \to \mathbb{R}^{n_l}, \quad l = 1, \dots, J, \quad \underline{\tilde{Q}}_J = \underline{I},$$

which satisfy

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|^2_{\mathbb{R}^{n_J}} \leq C_1^2 \frac{(r+1)^{2l}}{\lambda} \|\mathbf{u}\|^2_{\underline{A}} \quad \forall \mathbf{u} \in \mathbb{R}^{n_J},\ l = 1, \dots, J-1. \tag{4.9}$$

Keeping in mind the theoretical tools revisited in section 4.1, the proof of the weak approximation property (4.9) for the dGFEM is straightforward. We assume that the following geometric property holds for the agglomeration.

Assumption 2. *For each level $1 \leq l \leq J-1$ and given element agglomerates a_i^l, $i = 1, \dots, n_l$, there holds*

$$\operatorname{diam}(a_i^l) \leq C(r+1)^l h, \tag{4.10}$$

with r defined as in Theorem 4.5.

The linear mappings $\underline{\tilde{Q}}_l \colon \mathbb{R}^{n_J} \to \mathbb{R}^{n_l}$, $l = 1, \dots, J-1$, are then constructed as follows. We define for $u_h \in V_h^p$, represented by $u_h = \sum_{i=1}^{n_J} u_i \phi_i$, $\mathbf{u} = \{u_i\}_{i=1}^{n_J}$, the following:

$$\underline{\tilde{Q}}_k \mathbf{u} = \mathbf{w}^k, \quad \left(\mathbf{w}^k\right)_i = \frac{1}{\mu(a_i^l)} \int_{a_i^l} u_h \, d\mathbf{x}, \quad i = 1, \dots, n_k.$$

Then we have

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|_2^2 = \sum_{i=1}^{n_l} \|\mathbf{u} - \tilde{I}_l^J \mathbf{w}^l\|^2_{\ell^2\left(\mathcal{I}_i^{l,J}\right)} = \sum_{i=1}^{n_l} \|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|^2_{\ell^2\left(\mathcal{I}_i^{l,J}\right)}, \tag{4.11}$$

where for a set $T \subset \mathbb{N}^*$ the notation $\|\cdot\|_{\ell^2(T)}$ denotes the Euclidean norm restricted to the indices contained in T. In the continuous (broken) L^2-norm the expression can be rewritten as

$$\begin{aligned} \|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|^2_{\ell^2\left(\mathcal{I}_i^{l,J}\right)} &\leq Ch^{-d} \|u_h - \left(\mathbf{w}^l\right)_i\|^2_{0,a_i^l} \\ &\leq Ch^{-d} \operatorname{diam}\left(a_i^l\right)^2 \left(\sum_{\substack{\kappa \in \mathcal{T}_h, \\ \kappa \subset a_i^l}} |u_h|^2_{1,\kappa} + \sum_{\substack{e \in \mathcal{E}, \\ e \subset a_i^l}} h^{-1} \|[\![u_h]\!]\|^2_{0,e} \right), \end{aligned}$$

where in the last step we have applied the Poincaré–Friedrichs inequality (4.3). Finally, using assumption (4.10) and the coercivity of the bilinear form we get

$$\|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|^2_{\ell^2\left(\mathcal{I}_i^{l,J}\right)} \leq Ch^{-d} \left[C(r+1)^l h\right]^2 B_1(u_h, u_h) = Ch^{2-d}(r+1)^{2l} B_1(u_h, u_h).$$

Substituting into (4.11) together with (4.4) yields the result

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|_2^2 \leq (r+1)^{2l} \frac{C}{\varrho(\underline{A})} \|\mathbf{u}\|^2_{\underline{A}},$$

which gives the bound (4.9) and, consequently, the estimate (4.7).

4.3. Preconditioning of nonsymmetric discretizations. The convergence result developed in the previous section does not extend to the case of the nonsymmetric interior penalty discretization. Nevertheless, error estimates can be established for the NIPG as well if we take a GMRES iteration and a multilevel preconditioner for the symmetric part

$$B_0(u,v) := \int_\Omega \nabla_h u \cdot \nabla_h v \; d\mathbf{x} + \int_{\Gamma \cup \Gamma_{\text{int}}} [\![u]\!] \cdot [\![v]\!] \; ds,$$

corresponding to the stiffness matrix $\underline{A}_0 = \frac{1}{2}\left(\underline{A}_{NIPG} + \underline{A}_{NIPG}^T\right)$.

Let $c_1, c_2 > 0$ be defined as

$$c_1 = \lambda_{\min}\left(\underline{B}_{MG}\underline{A}_0\right), \quad c_2 = \|\underline{B}_{MG}\underline{A}_{\text{NIPG}}\|_{\underline{A}_0}.$$

Then, following the theory for minimal residual Krylov methods [12], the norm of the residual $\mathbf{r}^k$ after k steps of the GMRES method is bounded by

$$\|\mathbf{r}_k\|_{\underline{A}_0} \leq \left(1 - \frac{c_1^2}{c_2^2}\right)^{\frac{k}{2}} \|\mathbf{r}_0\|_{\underline{A}_0}.$$

Note that this result is valid for the GMRES method formulated with respect to the energy norm $\|\cdot\|_{\underline{A}_0}$, but it is of importance for the Euclidean inner product iteration as well, since there exist constants C_1, C_2 such that

$$C_1 h^d \|\mathbf{x}\|_2^2 \leq \|\mathbf{x}\|_{\underline{A}_0}^2 \leq C_2 h^{d-2} \|\mathbf{x}\|_2^2.$$

Thus the iteration count is increased only by a factor of order $\ln h$.

Usually, convergence bounds for preconditioners applied to nonsymmetric forms are obtained using a lemma by Schatz and a duality argument; cf. [36]. This technique cannot be applied to the NIPG bilinear form which is not adjoint consistent [3]. Hence, in the following we directly derive estimates for the constants c_1, c_2. First we need the following result.

LEMMA 4.6 (Antonietti and Ayuso [1]). *There exists a constant $\gamma > 0$ such that*

$$B_{NIPG}(v,w) \leq \gamma \left[B_0(v,v)\right]^{\frac{1}{2}} \left[B_0(w,w)\right]^{\frac{1}{2}} \quad \forall v \in V_h^p.$$

Now the following lemma is a consequence of Theorem 4.5.

LEMMA 4.7. *Let the nonsymmetric interior penalty discretization be solved with the GMRES algorithm. We further assume that the multilevel preconditioner described in section* 3 *is formulated for the symmetric part $\underline{A}_0$ of the stiffness matrix. Then there hold*

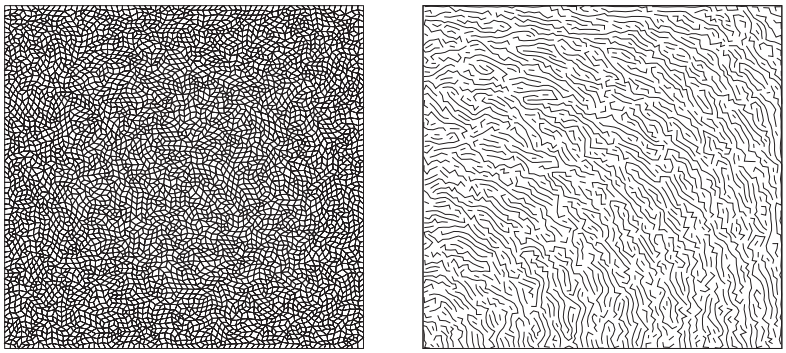
$$c_1 := \lambda_{\min}\left(\underline{B}_{MG}\underline{A}_0\right) \geq \frac{1}{C} \quad \textit{and} \quad c_2 := \|\underline{B}_{MG}\underline{A}_{NIPG}\|_{\underline{A}_0} \leq \gamma$$

with the constant C given in (4.8).

Proof. The lower bound follows directly from Theorem 4.5, since $\underline{A}_0$ is symmetric, and all assumptions hold due to the spectral equivalence between $B_{\text{SIPG}}(\cdot,\cdot)$ and $B_0(\cdot,\cdot)$; cf. Lemmas 4.3 and 4.6. For the upper estimate, we have

$$\begin{aligned}
\|\underline{B}_{MG}\underline{A}_{\text{NIPG}}\mathbf{x}\|_{\underline{A}_0}^2 &= (\underline{A}_0\underline{B}_{MG}\underline{A}_{\text{NIPG}}\mathbf{x}, \underline{B}_{MG}\underline{A}_{\text{NIPG}}\mathbf{x}) \\
&= (\underline{A}_{\text{NIPG}}\mathbf{x}, \underline{B}_{MG}\underline{A}_0\underline{B}_{MG}\underline{A}_{\text{NIPG}}\mathbf{x}) \leq \gamma \|(\underline{B}_{MG}\underline{A}_0\underline{B}_{MG})\underline{A}_{\text{NIPG}}\mathbf{x}\|_{\underline{A}_0}\|\mathbf{x}\|_{\underline{A}_0} \\
&\leq \gamma \|\underline{B}_{MG}\underline{A}_{\text{NIPG}}\mathbf{x}\|_{\underline{A}_0}\|\mathbf{x}\|_{\underline{A}_0},
\end{aligned}$$

where we used $\sigma(\underline{I} - \underline{B}_{MG}\underline{A}_0) \subset [0, 1 - \frac{1}{C}]$. Thus the norm of the operator is bounded by γ. ∎



(a) Example mesh, 5316 elements. (b) Constructed lines.

FIG. 5.1. *Result of the line creation heuristic. Advection vector* $\beta(\mathbf{x}) = (-x_2, x_1)^T$*; the lines formed by elements are represented by curves through the centers.*

5. Numerical examples. In this section we demonstrate the feasibility of the smoothed aggregation multigrid for the dGFEM. Numerical results have been obtained using the `PADGE` code [20] with the support of the `deal.II` libraries [5], Argonne's PETSc library, and the LAPACK linear algebra package.

5.1. Setting. With the exception of the test case including the NACA 0012 airfoil that is discussed in section 5.4, we restrict ourselves to the test domain $\Omega = (0,1)^2$. Two sequences of tesselations are considered for scalability experiments: The first sequence, $\{\mathcal{T}_i^s\}_{i=1}^5$, consists of nested quadrilateral meshes constructed by global uniform refinement of the unit square. The corresponding grid sizes are $h = \frac{1}{32}, \frac{1}{64}, \ldots, \frac{1}{512}$. The second mesh sequence, $\{\mathcal{T}_i^u\}_{i=1}^5$, is a collection of unstructured, isotropic, nonnested tesselations with $N = \{1335, 5316, 21104, 84269, 337075\}$ elements. The second-coarsest mesh, $\mathcal{T}_2^u$, is depicted in Figure 5.1(a).

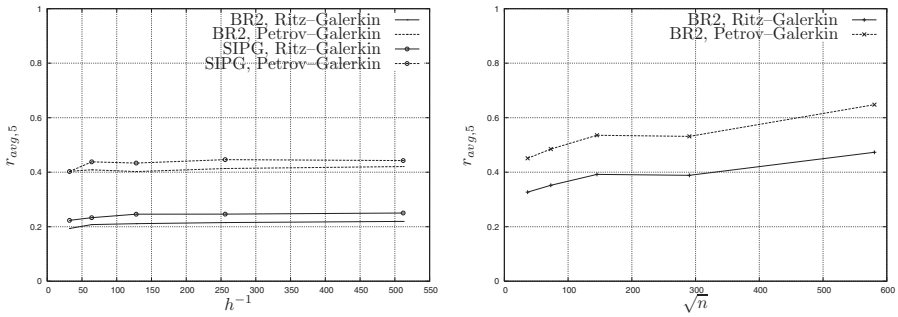
The algorithmic performance is measured by the average and the asymptotic residual reduction, denoted by r_{avg} and $r_{avg,k}$. After N iterations of the V-cycle evaluated in the Euclidean norm we have

$$r_{avg} := \left(\frac{\|\mathbf{r}^N\|_2}{\|\mathbf{r}^0\|_2}\right)^{\frac{1}{N}}, \quad r_{avg,k} := \left(\frac{\|\mathbf{r}^N\|_2}{\|\mathbf{r}^{N-k}\|_2}\right)^{\frac{1}{k}}, \quad k < N.$$

Additionally, in view of Theorem 4.5 we also approximate the reduction of the exact error in the energy norm $\|\underline{A}^{-1}\mathbf{b} - \mathbf{x}^N\|_{\underline{A}}$.

5.2. Poisson's equation. Since we are interested in the ideal scalability of the algorithm, we first consider a simple experiment with Poisson's equation having homogeneous Dirichlet boundary conditions and a constant source term $f \equiv 1$. For the linear algebraic system preconditioned by the smoothed aggregation algorithm, Theorem 4.5 predicts a condition number bounded by a low-order polynomial in terms of multigrid levels J, thus depending only slightly on the number of unknowns.

As a simple smoothing iteration for this problem we apply two pre- and post-smoothing steps of the pointwise, symmetric Gauß–Seidel method, abbreviated as V(2,2). The spectral radii $\varrho(\underline{A}_k)$, $k = 2, \ldots, J$, required for the construction of the smoothing polynomials, are approximated by a few steps of the Lanczos algorithm.



(a) uniform rectangular grid sequence $\{\mathcal{T}_i^s\}_{i=1}^5$. (b) unstructured grid sequence $\{\mathcal{T}_i^u\}_{i=1}^5$.

FIG. 5.2. *Asymptotic residual reduction $r_{avg,5}$ for the solution of Poisson's problem. V-cycle $V(2,2)$ with acceleration by the GMRES method.*

In all examples, a Krylov-type iterative method has been employed together with the V-cycle preconditioner. We used the so-called *flexible* variant [37] of the GMRES method minimizing the ℓ^2-norm of the residual. The number of levels varied from 5 to 9, such that the coarsest level consisted of fewer than 16 unknowns. Instead of a maximum number of iterations, a relative tolerance of $\|\mathbf{r}^k\|\|\mathbf{r}^0\|^{-1} \leq 10^{-10}$ for the Euclidean norm of the residual (coefficient) vector was chosen as an abort criterion, with the zero vector as initial approximation.

In Figure 5.2 the asymptotic residual reduction $r_{avg,5}$ is shown for the structured and unstructured mesh sequences. The values correspond to a piecewise bilinear SIPG discretization ($\delta := 10$) and the Bassi–Rebay scheme. They are plotted over the square root of elements, which is roughly the grid size h for the unstructured tesselations. Both the symmetric and the nonsymmetric multigrid variants scale optimally with h on the structured mesh sequence, and the average residual reduction seems to be bounded. Table 5.1 reflects the same behavior for the error reduction that is measured in the energy norm. Finally, the theoretical results are supported by the condition number estimates in Table 5.2, whereas the piecewise constant transfer operators cause the multilevel algorithm to deteriorate. The Ritz–Galerkin multigrid yields better results than the Petrov–Galerkin approach; however, we do not take into account the increased computational effort, which is caused by the higher density of the coarse level matrices. The additional presmoothing step for the Petrov–Galerkin multigrid has been omitted in the examples, as it did not seem to affect the convergence behavior of the method.

Results are less satisfactory for the unstructured mesh sequence; at least there is no clear evidence for an asymptotic behavior. This may be caused by the irregular shape of the agglomerates, which can already be seen in Figure 3.1. In general, AMG methods tend to give better results for element clusters with convex shape and approximately equal diameter; see, for example, the results presented in [40]. This corresponds to the theoretical Assumption 2, whose exact properties are difficult to verify in practice.

The multilevel method can be accelerated by a modified coarse grid correction

$$\mathbf{x}^l \leftarrow \mathbf{x}^l + \alpha I_{k-1}^k \mathbf{x}^{l-1}$$

with a factor $\alpha > 1$. The effect of this *overcorrection* is shown in the last row of

TABLE 5.1
Average error reduction with respect to the energy norm $\|\cdot\|_{\underline{A}}$. Results are shown for the structured (top) and unstructured mesh sequence. BR2 discretization, $V(2,2)$-cycle.

		$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
Point Gauß–Seidel	Ritz–Galerkin	0.189	0.195	0.199	0.201	0.198
	Petrov–Galerkin	0.373	0.385	0.393	0.395	0.399
	Ritz–Galerkin, $p=0$	0.005	0.007	0.008	0.006	0.021
Element Jacobi	Ritz–Galerkin	0.222	0.224	0.213	0.205	0.204
	Petrov–Galerkin	0.384	0.382	0.389	0.394	0.427
		$\mathcal{T}_1^u$	$\mathcal{T}_2^u$	$\mathcal{T}_3^u$	$\mathcal{T}_4^u$	$\mathcal{T}_5^u$
Point G. S.	Ritz–Galerkin	0.332	0.348	0.393	0.388	0.493
	Petrov–Galerkin	0.462	0.487	0.525	0.532	0.613
	overcorrection	0.321	0.331	0.362	0.408	0.506

TABLE 5.2
Condition number estimate $\operatorname{cond}(\underline{B}_{MG}\underline{A})$*, Ritz–Galerkin multigrid.*

	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
BR2, point G. S.	2.415	2.449	2.470	2.476	2.493
BR2, element Jacobi	3.051	3.058	3.053	3.046	3.040
SIPG, point G. S.	2.831	2.915	2.956	2.967	2.978
BR2, simple prolongation $\tilde{I}_{k-1}^k$	23.599	21.487	83.913	165.200	239.159

TABLE 5.3
Average and asymptotic residual reduction for the solution of Poisson's problem on a sequence of refined grids. NIPG discretization, Ritz–Galerkin multigrid with SOR smoother.

	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$	$\mathcal{T}_1^u$	$\mathcal{T}_2^u$	$\mathcal{T}_3^u$	$\mathcal{T}_4^u$	$\mathcal{T}_5^u$
$r_{avg,5}$	0.180	0.181	0.183	0.189	0.199	0.508	0.576	0.558	0.584	0.643
r_{avg}	0.211	0.215	0.223	0.232	0.237	0.544	0.608	0.595	0.623	0.683

Table 5.1 for the unstructured mesh sequence. Here, the experimentally determined scaling $\alpha = \frac{5}{3}$ of the coarse level operators yields an improvement of the average error reduction $\|\underline{A}^{-1}\mathbf{b} - \mathbf{x}^N\|_{\underline{A}}$ of about 0.1.

Further, the convergence result stated in section 4.3 for the nonsymmetric interior penalty discretization was verified for Poisson's equation. The experimental rates of convergence given in Table 5.3 are qualitatively similar to the results for the symmetric dGFEM variants.

Finally, we discuss the computational complexity of the method. Following the arguments in section 3.2, the dimensions n_i, $i = 1, \ldots, J$, of the level subspaces have to increase geometrically with i, which is determined by the agglomeration strategy. Further, the amount of work per cycle, except for the coarsest level, must behave like Cn_i. This yields an operation count like $\mathcal{O}(n_J)$ for the whole V-cycle. This is investigated in Table 5.4. It compares the nonzeros, denoted by #nnz, which are contained in the smoothed aggregation level matrices to the case of piecewise constant prolongation and restriction. The additional fill-in by the smoothed aggregation approach is significant, especially for piecewise constant finite elements, though much less severe for the Petrov–Galerkin approach than for the Ritz–Galerkin multigrid.

Timing results for Poisson's problem solved on the unstructured grid sequence are presented in Table 5.5. All runs were performed on an AMD Opteron 3.0 GHz system with 16 GB RAM (GNU C++). Both multigrid methods were tested together

TABLE 5.4
Ratio $\frac{1}{a_0}\sum_{k=1}^{J}\#nnz(\underline{A}_k)$ *for the smoothed aggregation coarse operators, BR2 discretization. By* a_0 *we denote the number of nonzeros when applying piecewise constant smoothing polynomials.*

$p=0$	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
Petrov–Galerkin	2.028	2.099	2.135	2.153	2.162
Ritz–Galerkin	3.797	4.310	4.719	5.016	5.233
$p=1$	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
Petrov–Galerkin	1.137	1.147	1.152	1.155	1.156
Ritz–Galerkin	1.372	1.443	1.499	1.540	1.569
$p=1$	$\mathcal{T}_1^u$	$\mathcal{T}_2^u$	$\mathcal{T}_3^u$	$\mathcal{T}_4^u$	$\mathcal{T}_5^u$
Petrov–Galerkin	1.160	1.172	1.180	1.187	1.169
Ritz–Galerkin	1.483	1.595	1.699	1.822	1.763

TABLE 5.5
Timing results for the unstructured grid sequence. The setup phase is split into (aggregation + matrix setup) and does not contain the assembly of the fine level matrix $\underline{A}_J$*. Measurements for a GMRES solver with ILU(0) preconditioner are given for comparison.*

Method	Card $\mathcal{T}_i^u$	Setup time	SOR smoother iter. time	Line Jacobi iter. time
Ritz–Galerkin multigrid	1335	0.07s + 0.12s	0.16s	0.52s
	5316	0.37s + 0.58s	0.85s	2.01s
	21104	2.87s + 3.39s	4.88s	9.47s
	84269	31s + 22s	23s	40s
	337075	366s + 135s	190s	278s
Petrov–Galerkin multigrid	1335	0.07s + 0.10s	0.20s	0.72s
	5316	0.37s + 0.39s	1.05s	2.80s
	21104	2.87s + 1.67s	5.34s	12.26s
	84269	30s + 7s	25s	53s
	337075	365s + 44s	165s	349s
GMRES with ILU	Iteration time: $\mathcal{T}_1^u$: 0.27s, $\mathcal{T}_2^u$: 4.42s, $\mathcal{T}_3^u$: 45s, $\mathcal{T}_4^u$: 464s, $\mathcal{T}_5^u$: 5110s			

with a symmetric SOR smoother and a forward x-line smoother in a V(2,2)-cycle. Results for an ILU(0) preconditioner are given for comparison. The table contains the setup time needed for aggregating the unknowns and assembling the coarse level matrices, but excludes the fine level matrix assembly which is performed for the ILU preconditioner, too.

In absolute numbers the current implementation is far from optimal. Especially the implementation of the Jones–Vassilevski algorithm does not scale properly with the grid size. For the solution of the isotropic scalar model problem the line-implicit smoother is overly expensive. However, even when considering the total computation time needed for the setup and solution phase, the multigrid algorithm is faster than the ILU preconditioned GMRES method especially on finer grids. Comparing the two multigrid methods, the Ritz–Galerkin needs fewer iterations, but the Petrov–Galerkin turns out to be still faster for the SOR smoother on the finest level. For the line-Jacobi test, the complicated smoothing iteration is clearly the most expensive operation. Hence the Petrov–Galerkin multigrid is slower than the Ritz–Galerkin multigrid; at least the difference seems to decrease with finer grids.

5.3. Linear convection-diffusion problem. We now return to the second-order differential equation (2.1) and consider the convection-diffusion problem, which is defined by the trivial diffusion matrix $\underline{a} \equiv \mu \underline{I}$ and the vector field $\beta(\mathbf{x}) = \gamma\,(-x_2, x_1)^T$.

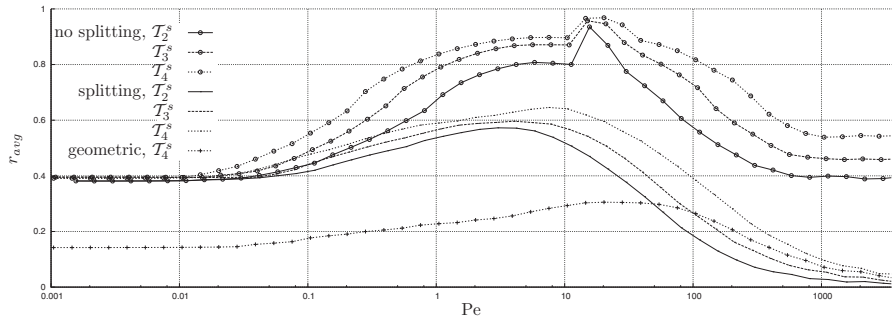


FIG. 5.3. *Péclet number study for Petrov–Galerkin variational multigrid. BR2 discretization, $V(1,1)$-cycle. Convergence study for varying Péclet number and comparison with geometric multigrid.*

We set $\mu = 10^{-3}$ and let $\gamma \in \mathbb{R}^+$ denote a variable scaling factor of the rigid rotation. The equation is complemented with Dirichlet boundary conditions for $\Gamma_{\mathrm{D}} = \{\mathbf{x} \in \Gamma\colon (x_1 = 1) \vee (x_2 = 0)\}$, while on the remaining part we apply a Neumann boundary condition, $\nabla u \cdot \mathbf{n} = g_{\mathrm{N}} \equiv 0$. The data g_{D} on the Dirichlet boundary Γ_{D} is given by

$$g_{\mathrm{D}}(x_1, x_2) := \begin{cases} 5(x_1 - 0.2) & \text{for} \quad 0.2 < x_1 \leq 0.4,\ x_2 = 0, \\ 1 & \text{for} \quad 0.4 < x_1 \leq 0.6,\ x_2 = 0, \\ 1 - 5(x_1 - 0.6) & \text{for} \quad 0.6 < x_1 \leq 0.8,\ x_2 = 0, \\ 0 & \text{elsewhere.} \end{cases}$$

We are interested in the behavior of the multigrid algorithm in the case of varying advection skew to the mesh. Therefore, the average residual reduction r_{avg} is computed over a range of values for the Péclet number $\mathrm{Pe} := \|\beta\|_\infty h \mu^{-1}$, where the maximum speed of advection occurring in the domain of interest is $\|\beta\|_\infty = \gamma$. The results shown in Figure 5.3 were computed for the structured mesh sequence $\{\mathcal{T}_i^s\}_{i=1}^5$ with bilinear finite elements. The algorithm was chosen to be a GMRES accelerated multigrid with a V(1,1) line-implicit Gauß–Seidel smoothing iteration.

We discuss the results for the Petrov–Galerkin multigrid. The set of the studied test cases can be divided into three different flow regimes; see Figure 5.3. In the first, approximately $\mathrm{Pe} \in [10^{-4}; 10^{-1}]$, the elliptic term stays predominant. As can be expected from the h-optimality tests in section 5.2 the average residual reduction is approximately the same for all meshes.

On the other hand, in the case of predominant convection, $\mathrm{Pe} > 100$, optimality of the algorithm is not fulfilled; i.e., the convergence rate increases with the number of unknowns. The behavior of the multigrid iteration improves significantly when employing a directionally implicit smoothing iteration, since the Gauß–Seidel method degrades into a (nearly) direct solver. The effect of the operator splitting $\underline{A} = \underline{A}^v + \underline{A}^c$ with a separate treatment of the convective and the diffusive operator, which was discussed in section 3.5, is visible in the flow regime $\mathrm{Pe} \in [10^{-1}; 100]$, where the diffusive and convective components balance each other. Compared to the smoothed aggregation algorithm without operator splitting, the average residual reduction is improved for all grid sizes.

As we are considering the sequence of nested, structured tesselations for this test problem, the smoothed aggregation multigrid can be compared directly to the classical geometric multigrid, where the hierarchy of levels stems from the previous refinement process. We choose the same problem parameters as before, but as interpolation operators I_k^{k-1}, I_{k-1}^k we use the injection $\iota\colon V_h^p \hookrightarrow V_H^p$, where $H := 2h$. The algorithm, which is analyzed in [16], uses one pre- and postsmoothing step of the line-Gauß–Seidel method. Results are shown in Figure 5.3. In the regime of predominant diffusion the geometric algorithm is superior to the algebraic approach. However, this changes with an increasing convection term, and the algorithm performs similarly to the AMG for (nearly) hyperbolic problems. Overall, we note that this comparison does not take into account the absolute computational effort, where the geometric multigrid is clearly superior due to its simpler algorithmic nature.

5.4. Linearized Navier–Stokes equations. As an example for nonlinear convection-diffusion, we consider a Newton multigrid approach for the Navier–Stokes equations. For laminar compressible flow the conservative state is given by $\mathbf{u} = [\rho, \rho\mathbf{v}, \rho E]^T$ with ρ denoting the density, $\mathbf{v}$ the velocity vector, and E the specific total energy. The equations are formulated in the Cartesian coordinate system as

$$\partial_t \mathbf{u} + \nabla \cdot (\mathcal{F}^c(\mathbf{u}) - \mathcal{F}^v(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{0} \quad \text{in } \Omega \subset \mathbb{R}^d \tag{5.1}$$

with convective and viscous fluxes $\mathcal{F}^c(\mathbf{u})$, $\mathcal{F}^v(\mathbf{u}, \nabla\mathbf{u})$ and complemented by suitable Dirichlet and Neumann boundary conditions; see, for example, the monograph by Feistauer, Felcman, and Straškraba [13] for details. The dG semilinear form is then derived similarly to the scalar case via reformulation as a first-order system and the substitution of stabilizing numerical fluxes [6, 21]. In our experiments the BR2 stabilization and bilinear finite elements were employed. For the numerical flux function we used the Vijayasundaram scheme [13].

Since we are interested in steady solutions, the variable $t > 0$ plays the role of a pseudotime variable, and the temporal accuracy is irrelevant. We linearize (5.1) with a semi-implicit Euler scheme; i.e., we solve

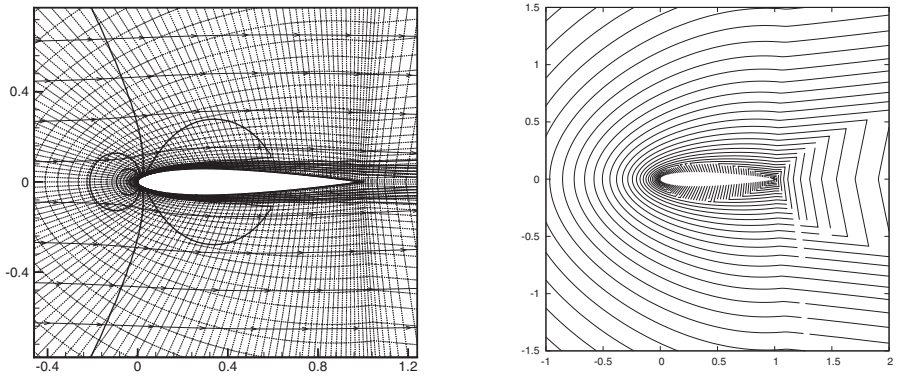
$$\left[\frac{1}{\Delta t}\underline{M} + \underline{J}(\mathbf{u}^n)\right] \mathbf{d} = -\nabla \cdot (\mathcal{F}^c(\mathbf{u}^n) - \mathcal{F}^v(\mathbf{u}^n, \nabla\mathbf{u}^n)) =: -\mathcal{N}(\mathbf{u}^n) \tag{5.2}$$

and set $\mathbf{u}^{n+1} = \mathbf{u}^n + \mathbf{d}$, where $\underline{J}(\mathbf{u}^n) := \frac{\partial \mathcal{N}}{\partial \mathbf{u}}(\mathbf{u}^n)$ denotes the Jacobian, $\underline{M}$ the mass matrix, and Δt is a suitable time step. We employed a local time step strategy controlled by a fixed global CFL number, $CFL = 10^2$, that is deactivated after six nonlinear steps.

The discrete linearized problem (5.2) exhibits a similar structure $\underline{J} = \underline{A}^v + \underline{A}^c$ as the convection-diffusion model case; however, it now corresponds to a vector-valued problem. Treating all degrees of freedom associated with a shape function as a block, the multigrid approach described in section 3 can be applied to the linearized Navier–Stokes equations. The recomputation of the interpolation and restriction operator in each step of the outer nonlinear iteration is avoided by constructing the smoothed transfer operators with respect to the discrete Laplacian [23].

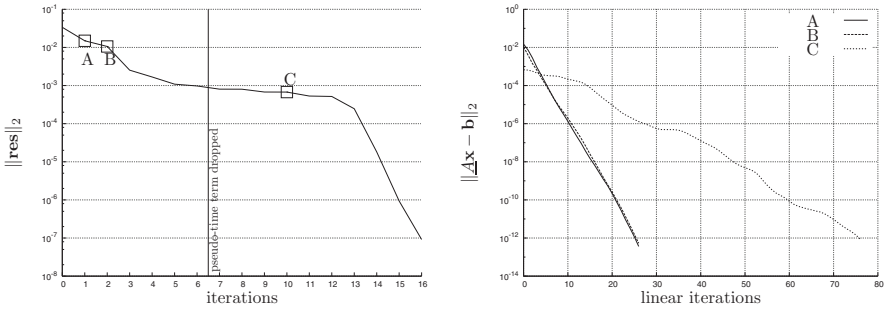
We restrict ourselves to a basic two-dimensional problem geometry. The test case under consideration is the symmetric NACA 0012 airfoil with an $\alpha = 0°$ angle of attack. We pose adiabatic no-slip boundary conditions at the wall and characteristic far-field boundary conditions. The flow is characterized by the quantities

$$\mathrm{Ma} = 0.5, \quad \mathrm{Re} = 5000,$$



(a) Solution: Mach number isolines. (b) Result of the line creation heuristic.

FIG. 5.4. *NACA* 0012 *test case (zoom).*



(a) Nonlinear convergence. (b) Linear convergence.

FIG. 5.5. *Smoothed aggregation multigrid as a preconditioner for the linearized compressible Navier–Stokes equations. Computational mesh with* 3072 *elements.*

which also determine the initial freestream state $\mathbf{u}^0$. The (structured) tesselations consist of 768, 3072, and 12288 quadrilaterals, respectively. Figure 5.4(a) shows a Mach number contour plot of the discrete solution.

As outlined in section 3.4 the coupling criterion of the line-implicit smoother is based on the momentum components $\rho\mathbf{v}$ of the flow state iterate $\mathbf{u}^k$. Figure 5.4(b) shows an illustration of the created lines, where the viscous effects have been well captured by the coupling criterion. To improve the stability of the smoothing iteration the line-implicit Gauß–Seidel scheme was extended to a multistage method. For this approach the smoothing iteration matrix $\underline{R}$ is replaced by a polynomial in $\underline{R}$ with optimized coefficients. Using the notation of an explicit Runge–Kutta method applied to the coefficient vector $\mathbf{x}^k$ we have

$$\mathbf{k}_0 = \mathbf{x}^k, \quad \mathbf{k}_j = \mathbf{x}^k + \alpha_j \underline{R}(\mathbf{b} - \underline{A}\mathbf{k}_{j-1}),\ j = 1, \ldots, s; \quad \mathbf{x}^k \leftarrow \mathbf{k}_s.$$

The coefficients $\{\alpha_j\}_{j=1,\ldots,s} := \{0.2075, 0.5915, 1\}$ were chosen according to [31].

Figure 5.5 shows some results for the Newton–Krylov algorithm preconditioned by a V(1,1)-cycle (Ritz–Galerkin). Three representative subproblems during the non-

TABLE 5.6
NACA 0012 *test case. Number of V-cycles required to reach a relative tolerance* $TOL < 10^{-10}$ *for Ritz–Galerkin multigrid, first linear subproblem.*

	768 elements	3072 elements	12288 elements
GMRES	22 ($r_{avg} = 0.346$)	25 (0.391)	27 (0.416)
BiCGStab	12	14	14

linear convergence process (cf. A, B, C in Figure 5.5(a)), have been plotted in Figure 5.5(b). The multilevel method performs significantly better in the presence of a pseudotime augmentation of the system matrix $\underline{J}(\mathbf{u})$. Considering the first nonlinear iteration, Table 5.6 lists the number of V-cycles required to reach a reduction of the residual norm by a factor of 10^{-10}. The applied Krylov-type iteration plays an important role, and the BiCGStab iteration requires about half of the steps needed for the GMRES method. For this method, however, the residual norm is no longer monotonically decreasing. Furthermore, due to the convective term and the simplifications made for the intergrid transfer operators, the convergence rate of the algorithm exhibits a slight dependence on the grid size h.

6. Conclusion. The concept of smoothed aggregation multigrid developed for the conforming finite element method on unstructured grids has been successfully applied to the discontinuous Galerkin discretization. The scalability of the algorithm has been analyzed theoretically, and results were supported by a number of numerical experiments. For the elliptic model problem the rate of convergence has been shown to be (nearly) independent of the number of unknowns. This could be observed for uniformly refined meshes and, to a reduced extent, also for general unstructured meshes. When considering equations of convection-diffusion type the intuitive nature of the aggregation multigrid lends itself to a combination with agglomeration approaches for the hyperbolic part of the problem operator. In this context, additional strategies from geometric multigrid, such as line-implicit smoothers and multistaging, were considered.

However, there are still some open questions. To the best of our knowledge, convergence theory for the smoothed aggregation approach is restricted only to the model case of elliptic problems. Moreover, the results are qualitative in nature; i.e., we typically obtain results with some generic constants. A proof of scalability and robustness for more general forms would be desirable. Apart from theory, the implementation of the smoothed aggregation multigrid exhibits some disadvantages. As for any global solution method, the algorithm is difficult to implement in parallel. Most important, the approach operates on linear systems containing considerable fill-in, even in the case of Petrov–Galerkin coarse level operators. This issue may become prohibitive in terms of memory; however, we hope it will be mitigated by future hardware developments.

Several problems that remain open will be the topic of our further work. The multigrid algorithm needs to be combined with a p-multilevel method or some related domain decomposition approach to yield a hierarchical solver for high-order numerical systems. Furthermore, the treatment of the convective terms needs improvement. This involves the formulation of the transfer operators as well as the development of better smoothing methods in order to deal with the linearized flow equations.

REFERENCES

[1] 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.

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

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

[4] O. Axelsson and P. Vassilevski, *Algebraic multilevel preconditioning methods.* I, Numer. Math., 56 (1989), pp. 157–177.

[5] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II—a general-purpose object-oriented finite element library*, ACM Trans. Math. Software, 33 (2007), article 24.

[6] F. Bassi, A. Crivellini, S. Rebay, and M. Savini, *Discontinuous Galerkin solution of the Reynolds-averaged Navier–Stokes and $k-\omega$ turbulence model equations*, Comput. & Fluids, 34 (2005), pp. 507–540.

[7] J. H. Bramble, *Multigrid Methods*, Pitman Res. Notes Math. Ser. 294, Longman Scientific & Technical, Harlow, UK, 1993.

[8] J. H. Bramble, J. E. Pasciak, J. P. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[9] S. C. Brenner and J. Zhao, *Convergence of multigrid algorithms for interior penalty methods*, Appl. Numer. Anal. Comput. Math., 2 (2005), pp. 3–18.

[10] T. Chan and W. Wan, *Robust multigrid methods for elliptic linear systems*, J. Comput. Appl. Math., 123 (1999), pp. 323–352.

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

[12] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[13] M. Feistauer, J. Felcman, and I. Straškraba, *Mathematical and Computational Methods for Compressible Flow*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, UK, 2003.

[14] 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.

[15] 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.

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

[17] H. Guillard, A. Janka, and P. Vaněk, *Analysis of an algebraic Petrov–Galerkin smoothed aggregation multigrid method*, Appl. Numer. Math., 58 (2006), pp. 1861–1874.

[18] H. Guillard and P. Vaněk, *An Aggregation Multigrid Solver for Convection-Diffusion Problems on Unstructured Meshes*, Technical report UCD-CCM-130, University of Colorado at Denver, Denver, 1998.

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

[20] R. Hartmann, J. Held, T. Leicht, and F. Prill, *PADGE, Parallel Adaptive Discontinuous Galerkin Environment*, Technical reference, DLR, Braunschweig, Germany, 2008.

[21] R. Hartmann and P. Houston, *Symmetric interior penalty DG methods for the compressible Navier–Stokes equations.* I*., Method fomulation*, Int. J. Numer. Anal. Model., 3 (2006), pp. 1–20.

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

[23] A. Janka, *Multigrid Methods for Compressible Laminar Flow*, Ph.D. thesis, INRIA, Nice, France, 2002.

[24] J. E. Jones and P. S. Vassilevski, *AMGe based on element agglomeration*, SIAM J. Sci. Comput., 23 (2001), pp. 109–133.

[25] J. Kraus and J. Synka, *An Agglomeration-Based Multilevel-Topology Concept with Application to 3D-FE Meshes*, Technical report, RICAM, Linz, Austria, 2004.

[26] J. K. Kraus and S. K. Tomar, *Multilevel preconditioning of two-dimensional elliptic problems discretized by a class of discontinuous Galerkin methods*, SIAM J. Sci. Comput., 30 (2008), pp. 684–706.
[27] C. Lasser and A. Toselli, *Overlapping preconditioners for discontinuous Galerkin approximations of second order problems*, in Proceedings of the Thirteenth International Conference on Domain Decomposition Methods, N. Debit, M. Garbey, R. Hoppe, J. Périaux, D. Keyes, and Y. Kuznetsov, eds., CIMNE, Barcelona, 2002.
[28] 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.
[29] D. J. Mavriplis, *Multigrid Techniques for Unstructured Meshes*, VKI Lecture Series VKI-LS 1995-02, Lecture Notes for 26th CFD Lecture Series of von Karman Institute, von Karman Institute for Fluid Dynamics, Rhode Saint Genese, Belgium, 1995.
[30] D. Mavriplis and V. Venkatakrishnan, *Agglomeration multigrid for two-dimensional viscous flows*, Comput. & Fluids, 24 (1995), pp. 553–570.
[31] T. O. Okusanya, D. L. Darmofal, and J. Peraire, *Algebraic multigrid for stabilized finite element discretizations of the Navier–Stokes equations*, Comput. Methods Appl. Mech. Engrg., 193 (2004), pp. 3667–3686.
[32] B. Philip and T. Chartier, *Adaptive Algebraic Smoothers*, Technical report LA-UR-07-6276, Los Alamos National Laboratory, Los Alamos, NM, 2008.
[33] S. Prudhomme, F. Pascal, J. T. Oden, and A. Romkes, *Review of A Priori Error Estimation for Discontinuous Galerkin Methods*, Technical report 00-27, TICAM, Austin, TX, 2000.
[34] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Ser. Comput. Math. 23, Springer-Verlag, Berlin, 1994.
[35] K. Shahbazi, *An explicit expression for the penalty parameter of the interior penalty method*, J. Comput. Phys., 205 (2005), pp. 401–407.
[36] A. Toselli and O. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer-Verlag, Berlin, 2005.
[37] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*, Cambridge Monogr. Appl. Comput. Math. 13, Cambridge University Press, Cambridge, UK, 2003.
[38] P. Vaněk, M. Brezina, and J. Mandel, *Convergence of algebraic multigrid based on smoothed aggregation*, Numer. Math., 88 (2001), pp. 559–579.
[39] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.
[40] M. Wabro, *AMGe—coarsening strategies and application to the Oseen equations*, SIAM J. Sci. Comput., 27 (2006), pp. 2077–2097.