DELFT UNIVERSITY OF TECHNOLOGY

REPORT 12-11

Scalable two-level preconditioning and deflation based on a piecewise constant subspace for (SIP)DG systems

P. van Slingerland and C. Vuik

ISSN 1389-6520

Reports of the Department of Applied Mathematical Analysis

Delft 2012

Scalable two-level preconditioning and deflation based on a piecewise constant subspace for (SIP)DG systems

P. van Slingerland C. Vuik

August 30, 2012

Abstract

This report is focused on the preconditioned Conjugate Gradient (CG) method for linear systems resulting from Symmetric Interior Penalty (discontinuous) Galerkin (SIPG) discretizations for stationary diffusion problems. In particular, it concerns two-level preconditioning strategies where the coarse space is based on piecewise constant DG basis functions. In this paper, we show that both the two-level preconditioner and the corresponding BNN (or ADEF2) deflation variant yield fast scalable convergence of the CG method (independent of the mesh element diameter). These theoretical results are illustrated by numerical experiments.

1 Introduction

The discontinuous Galerkin method can be interpreted as a finite volume method that uses piecewise polynomials of degree p rather than piecewise constant functions. As such, it combines the best of both classical finite element methods and finite volume methods, making it particularly suitable for handling non-matching grids and designing hp-refinement strategies. However, the resulting linear systems are usually larger than those for the aforementioned classical discretization schemes. This is due to the larger number of unknowns per mesh element. At the same time, the condition number typically increases with the number of mesh elements, the polynomial degree, and the stabilization factor [4, 18]. Problems with strong variations in the coefficients pose an extra challenge. Altogether, standard linear solvers often result in long computational times and/or low accuracy.

In search of efficient and scalable algorithms (for which the number of iterations does not increase with e.g. the number of mesh elements), much attention has been paid to subspace correction methods [23]. Examples include classical geometric (h-)multigrid [3, 10], spectral (p-)multigrid [9, 13], algebraic multigrid [14, 17], Schwarz domain decomposition [1, 8], and mixtures of these [2]. Usually, these methods can either be used as a standalone solver, or as a preconditioner in an iterative Krylov method. The latter tends to be more robust for problems with a few isolated 'bad' eigenvalues, as is typically the case for problems with large contrasts in the coefficients.

This research is focused on the preconditioned Conjugate Gradient (CG) method for linear systems resulting from Symmetric Interior Penalty (discontinuous) Galerkin (SIPG) discretizations for stationary diffusion problems. In particular, it concerns two-level preconditioning strategies where the coarse space is based on the piecewise constant DG basis functions.

The latter strategy has been introduced by Dobrev et al. [6]. In [19], their method has been carried over to deflation with the help of the analysis in [21]: it has been demonstrated numerically that both two-level variants yield fast and scalable CG convergence using (damped) block Jacobi smoothing.

To provide theoretical support for the latter, in this paper, we bound the condition number of the preconditioned/deflated system independently of the mesh element diameter. Such bounds are already available for the preconditioning variant with $p = 1$, as derived in [6] using the work of Falgout et al. [7]. Here, we extend these results by allowing $p \geq 1$. Furthermore, we

include BNN/ADEF2 deflation in the analysis. Additionally, we demonstrate that the required restrictions on the smoother are satisfied for (damped) block Jacobi smoothing. Finally, we extend the numerical support in [19] for these results by studying two test problems with strong variations in the coefficients.

The outline of this paper is as follows. Section 2 specifies both two-level methods for the linear SIPG systems under consideration. Section 3 derives an auxiliary regularity result, which is used to derive the main scalability result in Section 4. Numerical experiments are discussed in Section 5. Finally, we summarize the main conclusions in Section 6.

2 Methods & Assumptions

This section specifies the methods and assumptions that we consider in this paper. Section 2.1 specifies the diffusion model under consideration and discretizes it by means of the SIPG method. Section 2.2 discusses two two-level preconditioning strategies for solving the resulting linear system by means of the preconditioned CG method.

2.1 SIPG discretization for diffusion problems

Model problem We study the following diffusion problem on the d-dimensional domain Ω with source term f, scalar diffusion coefficient K (bounded below and above by positive constants), and a combination of Dirichlet and Neumann boundary conditions (specified by g_D on $\partial\Omega_D$ and g_N on $\partial\Omega_N$ with outward normal $\mathbf{n}$ respectively):

$$\begin{aligned} -\nabla\cdot(K\nabla u) &= f, && \text{in } \Omega, \\ u &= g_D, && \text{on } \partial\Omega_D, \\ K\nabla u\cdot\mathbf{n} &= g_N, && \text{on } \partial\Omega_N. \end{aligned} \tag{1}$$

We assume that the model parameters are chosen such that a weak solution of this model problem exists (cf. [15] for specific sufficient conditions). Furthermore, we assume that Ω is either an interval ($d=1$), polygon ($d=2$) or polyhedra ($d=3$).

Mesh To discretize the model problem (1), we subdivide Ω into mesh elements $E_1, ..., E_N$ with maximum element diameter h.

We assume that each mesh element E_i is affine-equivalent with a certain reference element E_0 that is an interval/polygon/polyhedra (independent of h) with mutually affine-equivalent edges. (2)

Note that all meshes consisting entirely of either intervals, triangles, tetrahedrons, parallelograms, or parallelepipeds satisfy this property.

Furthermore, we assume that the mesh is regular in the sense of [5, p. 124]. To specify this property, for all $i = 0, ..., N$, let h_i and ρ_i denote the diameter of E_i, and the diameter of the largest ball contained in E_i respectively. We can now define regularity as[1]:

$$\frac{h_i}{\rho_i} \lesssim 1, \qquad \forall i = 1, ..., N. \tag{3}$$

SIPG method Now that we have specified the mesh, we can construct an SIPG approximation for our model problem (1). To this end, define the test space V that contains each function that is a polynomial of degree p or lower within each mesh element, and that may be discontinuous at

[1] Throughout this paper, we use the symbol $\lesssim$ in expressions of the form "$F(x) \lesssim G(x)$ for all $x \in X$" to indicate that there exists a constant $C > 0$, independent of the variable x and the maximum mesh element diameter h (or the number of mesh elements), such that $F(x) \leq CG(x)$ for all $x \in X$. The symbol $\gtrsim$ is defined similarly.

the element boundaries. The SIPG approximation u_h is now defined as the unique element in this test space that satisfies the relation

$$B(u_h, v) = L(v), \qquad \text{for all test functions } v \in V, \tag{4}$$

where B and L are certain (bi)linear forms that are specified hereafter.

Details To specify these forms, we use the following notation: let Γ_h denote the collection of all edges $e = \partial E_i \cap \partial E_j$ in the interior. Similarly, let Γ_D and Γ_N denote the collection of all edges (points/lines/polygons) at the Dirichlet and Neumann boundary respectively. Additionally, for all edges e, let h_e denote the length of the largest mesh element adjacent to e for one-dimensional problems, the length of e for two-dimensional problems, and the square root of the surface area of e for three-dimensional problems. Finally, we introduce the usual trace operators for jumps and averages at the mesh element boundaries: in the interior, we define at $\partial E_i \cap \partial E_j$: $[\mathbf{v}] = \mathbf{v}_i \cdot \mathbf{n}_i + \mathbf{v}_j \cdot \mathbf{n}_j$, and $\{\mathbf{v}\} = \frac{1}{2}(\mathbf{v}_i + \mathbf{v}_j)$, where $\mathbf{v}_i$ denotes the trace of the (scalar or vector-valued) function $\mathbf{v}$ along the side of E_i with outward normal $\mathbf{n}_i$. Similarly, at the domain boundary, we define at $\partial E_i \cap \partial\Omega$: $[\mathbf{v}] = \mathbf{v}_i \cdot \mathbf{n}_i$, and $\{\mathbf{v}\} = \mathbf{v}_i$. Using this notation, the forms B and L in (4) can be defined as follows (for one-dimensional problems, the boundary integrals below should be interpreted as function evaluations of the integrand):

$$\begin{aligned}
B_\Omega(u_h, v) &= \sum_{i=1}^{N} \int_{E_i} K \nabla u_h \cdot \nabla v, \\
B_\sigma(u_h, v) &= \sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \frac{\sigma}{h_e} [u_h] \cdot [v], \\
B_r(u_h, v) &= -\sum_{e \in \Gamma_h \cup \Gamma_D} \int_e \Big(\{K \nabla u_h\} \cdot [v] + [u_h] \cdot \{K \nabla v\} \Big), \\
B(u_h, v) &= B_\Omega(u_h, v) + B_\sigma(u_h, v) + B_r(u_h, v), \\
L(v) &= \int_\Omega f v - \sum_{e \in \Gamma_D} \int_e \left([K \nabla v] + \frac{\sigma}{h_e} v \right) g_D + \sum_{e \in \Gamma_N} \int_e v g_N,
\end{aligned} \tag{5}$$

where σ is the so-called penalty parameter (discussed hereafter).

The SIPG method involves a penalty parameter σ, which penalizes the inter-element jumps to enforce weak continuity required for convergence. Its value may vary throughout the domain, and we assume that it is bounded below and above by positive constants (independent of the maximum element diameter h). Furthermore, we assume that the penalty parameter is sufficiently large so that the scheme is coercive [15, p. 38–40], i.e.:

$$B_\Omega(v, v) + B_\sigma(v, v) \lesssim B(v, v), \qquad \forall v \in V. \tag{6}$$

Linear system In order to compute the SIPG approximation defined by (4), it needs to be rewritten as a linear system. To this end, we choose a basis for the test space V: for each mesh element E_i, we set the basis function $\phi_1^{(i)}$ equal to zero in the entire domain, except in E_i, where it is equal to one. Similarly, we define higher-order basis functions $\phi_2^{(i)}, ..., \phi_m^{(i)}$, which are a higher-order polynomial in E_i and zero elsewhere. Note that, e.g., for one-dimensional problems, we have $m = p + 1$.

Details More specifically, the basis functions are constructed as follows. For all $i = 1, ..., N$, let $F_i : E_i \to E_0$ denote an invertible affine mapping (which exists by assumption (2)). Furthermore, let the functions $\phi_k^{(0)} : E_0 \to \mathbb{R}$ (with $k = 1, ..., m$) form a basis for the space of all polynomials of degree p and lower on the reference element (setting $\phi_1^{(0)} = 1$). Using this basis on the reference element, the basis function $\phi_k^{(i)}$ is zero in the entire domain, except in the mesh element E_i, where it reads $\phi_k^{(i)} = \phi_k^{(0)} \circ F_i$.

Now that we have defined the basis functions, we can express u_h as a linear combination of these functions:

$$u_h = \sum_{i=1}^{N} \sum_{k=1}^{m} u_k^{(i)} \phi_k^{(i)}. \tag{7}$$

The new unknowns $u_k^{(i)}$ in (7) can be determined by solving a linear system $A\mathbf{u} = \mathbf{b}$ of following form:

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & & \vdots \\ \vdots & & \ddots & \\ A_{N1} & \dots & & A_{NN} \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_N \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_N \end{bmatrix}, \tag{8}$$

where the blocks all have dimension m, and where, for all $i, j = 1, ..., N$:

$$A_{ji} = \begin{bmatrix} B(\phi_1^{(i)}, \phi_1^{(j)}) & B(\phi_2^{(i)}, \phi_1^{(j)}) & \dots & B(\phi_m^{(i)}, \phi_1^{(j)}) \\ B(\phi_1^{(i)}, \phi_2^{(j)}) & B(\phi_2^{(i)}, \phi_2^{(j)}) & & \vdots \\ \vdots & & \ddots & \\ B(\phi_1^{(i)}, \phi_m^{(j)}) & \dots & & B(\phi_m^{(i)}, \phi_m^{(j)}) \end{bmatrix}, \quad \mathbf{u}_i = \begin{bmatrix} u_1^{(i)} \\ u_2^{(i)} \\ \vdots \\ u_m^{(i)} \end{bmatrix}, \quad \mathbf{b}_j = \begin{bmatrix} L(\phi_1^{(j)}) \\ L(\phi_2^{(j)}) \\ \vdots \\ L(\phi_m^{(j)}) \end{bmatrix}. \tag{9}$$

This system is obtained by substituting the expression (7) for u_h and the basis functions $\phi_\ell^{(j)}$ for v into (4). Note that A is Symmetric and Positive-Definite (SPD), as the bilinear form B is SPD (cf. (5) and (6)). Once the unknowns $u_k^{(i)}$ are solved from the system, the final SIPG approximation u_h can be obtained from (7).

2.2 Two-level preconditioning and deflation

In the previous section, we obtained a linear SIPG system of the form $A\mathbf{x} = \mathbf{b}$, where A is SPD. To solve this system, we consider the preconditioned CG method. In particular, we focus on a two-level preconditioner introduced by Dobrev et al. [6], and the corresponding BNN-deflation variant. In this section, both variants are discussed.

Preconditioning variant The two-level preconditioner is defined in terms of a coarse correction operator $Q \approx A^{-1}$ that switches from the original DG test space to a coarse subspace, then performs a correction that is now simple in this coarse space, and finally switches back to the original DG test space. In this paper, we study the coarse subspace that consists of all piecewise constant functions.

More specifically, the coarse correction operator Q is defined as follows. Let R denote the so-called restriction operator such that $A_0 := RAR^T$ is the SIPG matrix for polynomial degree $p = 0$. In other words, R is the following matrix with blocks of size $1 \times m$:

$$R = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1N} \\ R_{21} & R_{22} & & \vdots \\ \vdots & & \ddots & \\ R_{N1} & \dots & & R_{NN} \end{bmatrix}, \qquad R_{ii} = \begin{bmatrix} 1 & 0 \dots & 0 \end{bmatrix}, \qquad R_{ij} = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix},$$

for all $i, j = 1, ..., N$ and $i \neq j$. Using this notation, the coarse correction operator is defined as $Q := R^T A_0^{-1} R$.

The two-level preconditioner combines this operator with a nonsingular smoother $M_{\text{prec}}^{-1} \approx A^{-1}$ with the property

$$M_{\text{prec}} + M_{\text{prec}}^T - A \text{ is SPD.} \tag{10}$$

It can be seen that this requirement is satisfied for (block) Gauss-Seidel smoothing (in that case, $M_{\text{prec}} + M_{\text{prec}}^T - A$ is the (block) diagonal of A). Furthermore, it will be shown in Section 4.4 that

this requirement is satisfied for block Jacobi smoothing. The result $\mathbf{y} = P_{\text{prec}}^{-1}\mathbf{r}$ of applying the two-level preconditioner to a vector $\mathbf{r}$ can now be computed in three steps:

$$\begin{aligned} \mathbf{y}^{(1)} &= M_{\text{prec}}^{-1}\mathbf{r} && \text{(pre-smoothing)}, \\ \mathbf{y}^{(2)} &= \mathbf{y}^{(1)} + Q(\mathbf{r} - A\mathbf{y}^{(1)}) && \text{(coarse correction)}, \\ \mathbf{y} &= \mathbf{y}^{(2)} + M_{\text{prec}}^{-T}(\mathbf{r} - A\mathbf{y}^{(2)}) && \text{(post-smoothing)}. \end{aligned} \tag{11}$$

The operator P_{prec} is SPD assuming that (10) is satisfied [22, p. 66].

BNN Deflation variant Basically, the BNN deflation variant is obtained by turning (11) inside out, and using an SPD smoother $M_{\text{defl}}^{-1} \approx A^{-1}$ (such as block Jacobi). We do not impose a condition of the form (10) at this point. The result $\mathbf{y} = P_{\text{defl}}^{-1}\mathbf{r}$ of applying the BNN deflation technique to a vector $\mathbf{r}$ can now be computed as:

$$\begin{aligned} \mathbf{y}^{(1)} &:= Q\mathbf{r} && \text{(pre-coarse correction)}. \\ \mathbf{y}^{(2)} &:= \mathbf{y}^{(1)} + M_{\text{defl}}^{-1}(\mathbf{r} - A\mathbf{y}^{(1)}) && \text{(smoothing)}, \\ \mathbf{y} &:= \mathbf{y}^{(2)} + Q(\mathbf{r} - A\mathbf{y}^{(2)}) && \text{(post-coarse correction)}. \end{aligned} \tag{12}$$

The operator P_{defl} is SPD for any SPD smoother M_{defl}^{-1}, as can be shown using the more abstract analysis in [22, p. 66].

Finally, we stress that the BNN deflation variant can be implemented more efficiently in a CG algorithm by using the so-called ADEF2 deflation variant. The latter is obtained by skipping the pre-coarse correction step in (12). ADEF2 yields the same iterates as BNN, as long as the starting vector $\mathbf{x}_0$ is pre-processed according to: $\mathbf{x}_0 \mapsto Q\mathbf{b} + (I - AQ)^T\mathbf{x}_0$ [21]. Furthermore, ADEF2 requires only 2 mat-vecs and 1 smoothing step per iteration, whereas the preconditioning variant (11) requires 3 mat-vecs and 2 smoothing steps. In Section 5, we will compare the overall numerical efficiency of these two methods. However, for the theoretical purposes in this paper, it is more convenient to study BNN than ADEF2.

Additional smoother requirements To derive the theoretical results in this paper, we require additional assumptions on the smoothers. To specify these, for any SPD matrix M, let $\pi_M := R^T(RMR^T)^{-1}RM$ denote the projection onto the coarse space $\text{Range}(R^T)$ that yields the best approximation in the M-norm (cf. [7]). Additionally, for any nonsingular matrix M such that $M + M^T - A$ is SPD, define the symmetrization $\widetilde{M} := M^T(M + M^T - A)^{-1}M$.

Using this notation, we can now specify the additional smoother requirements:

$$2M_{\text{defl}} - A \text{ is SPD}, \tag{13}$$

$$h^{2-d}\mathbf{v}^T\widetilde{M}_{\text{prec}}\mathbf{v} \lesssim \mathbf{v}^T\mathbf{v}, \qquad \forall \mathbf{v} \in \text{Range}(I - \pi_I), \tag{14}$$

$$h^{2-d}\mathbf{v}^T M_{\text{defl}}\mathbf{v} \lesssim \mathbf{v}^T\mathbf{v}, \qquad \forall \mathbf{v} \in \text{Range}(I - \pi_I). \tag{15}$$

It will be shown in Section 4.4 that these requirements are satisfied for (damped) block Jacobi smoothing (a similar strategy can be used to show (14) for block Gauss-Seidel smoothing with blocks of size $m \times m$).

The conditions (10) and (13) imply that "the smoother iteration is a contraction in the A-norm" [7, p. 473]. The main idea behind the conditions (14) and (15) is that the smoother should scale with h^{2-d} in the same way that A does, and that $\widetilde{M}$ is an efficient preconditioner for A in the space orthogonal to the coarse space $\text{Range}(R^T)$ [22, p. 78]. A slightly stronger version of (14) is also used in [6] to establish scalable convergence.

3 Auxiliary result: regularity on the diagonal of A

We want to show that the linear solvers discussed in the previous section are scalable (independent of h). To this end, we require an auxiliary result that roughly states that the diagonal blocks of

A all behave in a similar manner in the space orthogonal to the coarse space. This section derives this result using the mesh assumptions (2) and (3). In Section 4, we will apply this result to obtain the desired scalability of both two-level methods. The final result of this section is given in Theorem 3.3 in Section 3.3. To show this result, we require two intermediate results, which are discussed in Section 3.1 and Section 3.2 respectively.

3.1 Intermediate result I: using regularity

The first intermediate result is a rather abstract property of the mesh. To state this result, recall the mapping $F_i : E_i \to E_0$ (cf. Section 2). Because this mapping is invertible and affine by assumption (2), there exists an invertible matrix $G_i \in \mathbb{R}^{d\times d}$ and a vector $\mathbf{g}_i \in \mathbb{R}^d$ such that $F_i(\mathbf{x}) = G_i\mathbf{x} + \mathbf{g}_i$ for all $\mathbf{x} \in E_i$. Next, let $|G_i^{-1}|$ denote the determinant of G_i^{-1}, and define $Z_i := |G_i^{-1}|G_i^T G_i$. We now have the following result[2]:

Lemma 3.1 *The matrix Z_i above satisfies the following relation:*

$$1 \lesssim \lambda_{\min}(h^{2-d}Z_i) \leq \lambda_{\max}(h^{2-d}Z_i) \lesssim 1, \qquad \forall i = 1, ..., N. \tag{16}$$

To show this result, we use the regularity of the mesh (3). Furthermore, we use the following relations [5, p. 120–122][3]:

$$|G_i^{-1}| = \frac{\text{meas}(E_i)}{\text{meas}(E_0)}, \qquad \|G_i\|_2 \leq \frac{h_0}{\rho_i}, \qquad \|G_i^{-1}\|_2 \leq \frac{h_i}{\rho_0}. \tag{17}$$

We can now show Lemma 3.1:

Proof (of Lemma 3.1) Because $Z_i := |G_i^{-1}|G_i^T G_i$, and G_i is invertible, we have (cf. [16, p. 26]):

$$\lambda_{\max}(h^{2-d}Z_i) = h^{2-d}|G_i^{-1}|\lambda_{\max}(G_i^T G_i) = h^{2-d}|G_i^{-1}|\|G_i\|_2^2,$$
$$\lambda_{\min}(h^{2-d}Z_i) = h^{2-d}|G_i^{-1}|\frac{1}{\lambda_{\max}\left((G_i^T G_i)^{-1}\right)} = h^{2-d}|G_i^{-1}|\frac{1}{\|G_i^{-1}\|_2^2}.$$

Applying the relations in (17), using that meas(E_0), h_0 and ρ_0 do not depend on h, and observing that $\rho_i^d \lesssim \text{meas}(E_i) \lesssim h_i^d$ (for all $i = 1, ..., N$), we may write:

$$\lambda_{\max}(h^{2-d}Z_i) \leq h^{2-d}\frac{\text{meas}(E_i)}{\text{meas}(E_0)}\left(\frac{h_0}{\rho_i}\right)^2 \lesssim \frac{\text{meas}(E_i)}{h^d}\left(\frac{h}{\rho_i}\right)^2 \lesssim \left(\frac{h_i}{h}\right)^d\left(\frac{h}{\rho_i}\right)^2,$$
$$\lambda_{\min}(h^{2-d}Z_i) \geq h^{2-d}\frac{\text{meas}(E_i)}{\text{meas}(E_0)}\left(\frac{\rho_0}{h_i}\right)^2 \gtrsim \frac{\text{meas}(E_i)}{h^d}\left(\frac{h}{h_i}\right)^2 \gtrsim \left(\frac{\rho_i}{h}\right)^d\left(\frac{h}{h_i}\right)^2.$$

Hence, the proof is completed if we can show that

$$1 \leq \frac{h}{h_i} \leq \frac{h}{\rho_i} \lesssim 1, \qquad \forall i = 1, ..., N.$$

The first two inequalities follow from the fact that $\rho_i \leq h_i \leq h$. The last inequality follows as a special case from (3). Hence, we obtain (16), which completes the proof. ■

3.2 Intermediate result II: the desired result in terms of 'small' bilinear forms

The second intermediate result concerns the individual diagonal blocks of A in terms of 'small' bilinear forms. To state this result, we require the following notation: let V_0 denote the space of all polynomials of degree p and lower defined on the reference element E_0. Additionally, let Γ_i denote the set of all edges of E_i that are either in the interior or at the Dirichlet boundary. Furthermore,

[2] Throughout this paper, $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalue of a matrix with real eigenvalues respectively.

[3] Throughout this paper, meas(.) denotes the Lesbesque measure.

let Γ_0 denote the set of all edges of the reference element E_0. Next, define the following bilinear forms[4]:

$$B_\Omega^{(i)}(v,w) = \int_{E_i} K\nabla\left(v\circ F_i\right)\cdot\nabla\left(w\circ F_i\right), \qquad B_\Omega^{(0)}(v,w) = \int_{E_0}\nabla v\cdot\nabla w,$$
$$B_\sigma^{(i)}(v,w) = \sum_{e\in\Gamma_i}\int_e \frac{\sigma}{h_e}\left[v\circ F_i\right]\cdot\left[w\circ F_i\right], \qquad B_\sigma^{(0)}(v,w) = \sum_{e\in\Gamma_0}\int_e [v]\cdot[w],$$

for all $v, w \in V_0$ and $i = 1, ..., N$. Using this notation, we now have the following result:

Lemma 3.2 *The bilinear forms above satisfy the following relations:*

$$B_\Omega^{(0)}(w,w) \lesssim h^{2-d}B_\Omega^{(i)}(w,w) \lesssim B_\Omega^{(0)}(w,w), \qquad \forall w\in V_0, \qquad \forall i = 1,...,N. \tag{18}$$
$$0 \le h^{2-d}B_\sigma^{(i)}(w,w) \lesssim B_\sigma^{(0)}(w,w), \qquad \forall w\in V_0, \qquad \forall i = 1,...,N. \tag{19}$$

To show this result, we use the mesh properties (2) and (3), and the assumption that the diffusion coefficient and the penalty parameter are bounded above and below by positive constants (independent of h). We discuss the proof in four parts hereafter.

Proof (of (18) in Lemma 3.2) Because the diffusion coefficient K is bounded below and above by positive constants (independent of h), we may write (all displayed relations below are for all $w \in V_0$ and for all $i = 1, ..., N$):

$$\int_{E_i}\nabla\left(w\circ F_i\right)\cdot\nabla\left(w\circ F_i\right) \lesssim B_\Omega^{(i)}(w,w) \lesssim \int_{E_i}\nabla\left(w\circ F_i\right)\cdot\nabla\left(w\circ F_i\right). \tag{20}$$

Next, we apply the chain rule, using that the Jacobian of F_i is equal to G_i:

$$\int_{E_i}\nabla\left(w\circ F_i\right)\cdot\nabla\left(w\circ F_i\right) = \int_{E_i}\Big(\big(G_i\nabla w\big)\circ F_i\Big)\cdot\Big(\big(G_i\nabla w\big)\circ F_i\Big).$$

A change of variables (from $\mathbf{x}\in E_i$ to $F_i(\mathbf{x})\in E_0$) introduces a factor $|G_i^{-1}|$:

$$\int_{E_i}\nabla\left(w\circ F_i\right)\cdot\nabla\left(w\circ F_i\right) = \int_{E_0}|G_i^{-1}|\big(G_i\nabla w\big)\cdot\big(G_i\nabla w\big) = \int_{E_0}(\nabla w)^T\underbrace{|G_i^{-1}|G_i^TG_i}_{=Z_i}\nabla w.$$

Substitution of this expression into (20) and multiplication with h^{2-d} yields:

$$\int_{E_0}(\nabla w)^T(h^{2-d}Z_i)(\nabla w) \lesssim h^{2-d}B_\Omega^{(i)}(w,w) \lesssim \int_{E_0}(\nabla w)^T(h^{2-d}Z_i)(\nabla w).$$

Application of Lemma 3.1 gives:

$$\underbrace{\int_{E_0}\nabla w\cdot\nabla w}_{=B_\Omega^{(0)}(w,w)} \lesssim h^{2-d}B_\Omega^{(i)}(w,w) \lesssim \underbrace{\int_{E_0}\nabla w\cdot\nabla w}_{=B_\Omega^{(0)}(w,w)},$$

which completes the proof of (18). ■

Proof (of (19) in Lemma 3.2 for 1D problems) Because the penalty parameter σ is bounded below and above by positive constants (independent of h), and because $B_\sigma^{(i)}$ is SPSD, it follows that (all displayed relations below are for all $w \in V_0$ and for all $i = 1, ..., N$):

$$0 \le h^{2-d}B_\sigma^{(i)}(w,w) \lesssim \sum_{e\in\Gamma_i}\int_e\frac{h^{2-d}}{h_e}\left[w\circ F_i\right]\cdot\left[w\circ F_i\right]. \tag{21}$$

For one-dimensional problems, the integral over an edge e should be interpreted as the evaluation of the integrand. Furthermore, h_e denotes the size of the largest mesh element adjacent to e, so regularity (3) implies that $\frac{h}{h_e}\lesssim 1$ for all e. Hence, we may write (for $d = 1$):

$$0 \le hB_\sigma^{(i)}(w,w) \lesssim \sum_{e\in\Gamma_i}\left[w\circ F_i(e)\right]\cdot\left[w\circ F_i(e)\right].$$

[4]Here, the trace operators are defined as before by extending the function to be zero outside E_0 and E_i.

Finally, observe that, for all $e \in \Gamma_i$, the transformed edge value $F_i(e) =: e_0 \in \Gamma_0$. Different $e \in \Gamma_i$ yield different $e_0 \in \Gamma_0$, although not all $e_0 \in \Gamma_0$ are reached in the presence of Neumann boundary conditions. Nevertheless, we may write:

$$0 \leq hB_\sigma^{(i)}(w,w) \lesssim \underbrace{\sum_{e_0 \in \Gamma_0} [w(e_0)] \cdot [w(e_0)]}_{=B_\sigma^{(0)}(w,w)}.$$

This completes the proof of (19) for one-dimensional problems. ■

Proof (of (19) in Lemma 3.2 for 2D problems) Similar to the one-dimensional case, we have (21). For two-dimensional problems, the edges e are line segments and $h_e = \text{meas}(e)$, i.e. the length of e. Hence, for all e there exists an invertible affine mapping $r_e : [0,1] \to e$. By definition of the line integral over e, we may now rewrite (21) as (using $d = 2$):

$$0 \leq B_\sigma^{(i)}(w,w) \lesssim \sum_{e \in \Gamma_i} \frac{1}{h_e} \int_0^1 [w \circ F_i \circ r_e(t)] \cdot [w \circ F_i \circ r_e(t)] \, |r_e'| \, \mathrm{d}t$$

Because $r_e(t)$ is affine, its derivative r_e' is a constant and:

$$|r_e'| = \int_0^1 |r_e'| \, \mathrm{d}t = \int_e 1 \, \mathrm{d}e = \text{meas}(e) = h_e.$$

Hence:

$$0 \leq B_\sigma^{(i)}(w,w) \lesssim \sum_{e \in \Gamma_i} \int_0^1 [w \circ F_i \circ r_e(t)] \cdot [w \circ F_i \circ r_e(t)] \, \mathrm{d}t.$$

Next, consider a single $e \in \Gamma_i$: note that $F_i \circ r_e([0,1]) = F_i(e) =: e_0 \subset \partial E_0$, and define the invertible affine mapping $r_{e_0} := F_i \circ r_e$. As above, we have that $|r_{e_0}'| = \text{meas}(e_0)$. By definition of the line integral over e_0, we may now write (using that $\text{meas}(e_0)$ does not depend on h):

$$\int_0^1 [w \circ F_i \circ r_e(t)] \cdot [w \circ F_i \circ r_e(t)] \, \mathrm{d}t = \frac{1}{\text{meas}(e_0)} \int_{e_0} [w] \cdot [w] \lesssim \int_{e_0} [w] \cdot [w].$$

Next, we apply this strategy for all $e \in \Gamma_i$, which yield different (disjunct) $e_0 \subset \partial E_0$, although the entire boundary of E_0 is not reached in the presence of Neumann boundary conditions. Combining the results, we may write (after possible repartitioning of the edges e_0):

$$0 \leq B_\sigma^{(i)}(w,w) \lesssim \underbrace{\sum_{e_0 \in \Gamma_0} \int_{e_0} [w] \cdot [w]}_{=B_\sigma^{(0)}(w,w)}.$$

This completes the proof of (19) for two-dimensional problems ($d = 2$). ■

Proof (of (19) in Lemma 3.2 for 3D problems) The proof is similar to the two-dimensional case, except that we are now dealing with surface integrals rather than line integrals. Similar the one-dimensional case, we have (21). For three-dimensional problems, the faces e are polygons and $h_e = \sqrt{\text{meas}(e)}$, i.e. the square root of the surface area of e. Because all faces are mutually affine-equivalent (2), for all e, there exists an invertible affine mapping $r_e : D \to e$ for some polygon $D \subset \mathbb{R}^2$ (independent of h). By definition of the surface integral over e, we may rewrite (21) as (using $d = 3$):

$$0 \leq h^{-1} B_\sigma^{(i)}(w,w) \lesssim \sum_{e \in \Gamma_i} \frac{1}{h \, h_e} \int_D [w \circ F_i \circ r_e(u,v)] \cdot [w \circ F_i \circ r_e(u,v)] \left| \frac{\partial r_e}{\partial u} \times \frac{\partial r_e}{\partial v} \right| \mathrm{d}u \, \mathrm{d}v.$$

Because $r_e(u,v)$ is affine, it follows that its derivatives $\frac{\partial r}{\partial u}$ and $\frac{\partial r}{\partial v}$ are constant, and:

$$\left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| = \frac{1}{\text{meas}(D)} \int_D \left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| \mathrm{d}u \, \mathrm{d}v = \frac{1}{\text{meas}(D)} \int_e 1 \, \mathrm{d}e = \frac{\text{meas}(e)}{\text{meas}(D)} = \frac{h_e^2}{\text{meas}(D)},$$

Hence:

$$0 \leq h^{-1} B_\sigma^{(i)}(w,w) \lesssim \sum_{e \in \Gamma_i} \frac{h_e}{h \, \text{meas}(D)} \int_D [w \circ F_i \circ r_e(u,v)] \cdot [w \circ F_i \circ r_e(u,v)] \, \mathrm{d}u \, \mathrm{d}v$$

Because e can be contained in a circle with diameter h, it follows that $\text{meas}(e) \lesssim h^2$, hence $\frac{h_e}{h} \lesssim 1$:

$$0 \leq h^{-1} B_\sigma^{(i)}(w,w) \lesssim \sum_{e \in \Gamma_i} \frac{1}{\text{meas}(D)} \int_D [w \circ F_i \circ r_e(u,v)] \cdot [w \circ F_i \circ r_e(u,v)] \, \mathrm{d}u \, \mathrm{d}v$$

Next, consider a single $e \in \Gamma_i$: note that $F_i \circ r_e(D) = F_i(e) =: e_0 \subset \partial E_0$, and define the invertible affine mapping $r_{e_0} := F_i \circ r_e$. As above, we have that $\left|\frac{\partial r_{e_0}}{\partial u} \times \frac{\partial r_{e_0}}{\partial v}\right| = \frac{\mathrm{meas}(e_0)}{\mathrm{meas}(D)}$. By definition of the surface integral over e_0, we may now write (using that $\mathrm{meas}(e_0)$ does not depend on h):

$$\frac{1}{\mathrm{meas}(D)} \int_D [w \circ F_i \circ r_e(u,v)] \cdot [w \circ F_i \circ r_e(u,v)] \,\mathrm{d}u\,\mathrm{d}v = \frac{1}{\mathrm{meas}(e_0)} \int_{e_0} [w] \cdot [w] \lesssim \int_{e_0} [w] \cdot [w] \,.$$

Next, we apply this strategy for all $e \in \Gamma_i$, which yield different (disjunct) $e_0 \subset \partial E_0$, although the entire boundary of E_0 is not reached in the presence of Neumann boundary conditions. Combining the results, we may write (after possible repartitioning of the edges e_0):

$$0 \le h^{-1} B_\sigma^{(i)}(w,w) \lesssim \underbrace{\sum_{e_0 \in \Gamma_0} \int_{e_0} [w] \cdot [w] \,.}_{= B_\sigma^{(0)}(w,w)}$$

This completes the proof of (19) for three-dimensional problems ($d = 3$). ■

3.3 Final auxiliary result: regularity on the diagonal of A

Using the intermediate results in the previous sections, we can now show the final result of this section, which roughly states that the diagonal blocks of A all behave in a similar manner in the space orthogonal to the coarse space. To state this result, we require the following notation: suppose that A_Ω results from the bilinear form $h^{2-d}B_\Omega$ in the same way that A results from the bilinear form B: this is established by substituting A_Ω for A and $h^{2-d}B_\Omega$ for B in (8) and (9). Similarly, suppose that the matrices A_σ and A_r result from the bilinear forms $h^{2-d}B_\sigma$ and $h^{2-d}B_r$ respectively. Altogether, we may write: $A = h^{d-2}(A_\Omega + A_\sigma + A_r)$. Finally, let D_σ be the result of extracting the diagonal blocks of size $m \times m$ from A_σ. Using this notation, we now have the following result:

Theorem 3.3 *The matrices A_Ω and D_σ above satisfy the following relations:*

$$\mathbf{v}^T\mathbf{v} \lesssim \mathbf{v}^T A_\Omega \mathbf{v}, \qquad \forall \mathbf{v} \in \mathrm{Range}(I - \pi_I), \tag{22}$$

$$\mathbf{v}^T A_\Omega \mathbf{v} \lesssim \mathbf{v}^T\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{mN}, \tag{23}$$

$$0 \le \mathbf{v}^T D_\sigma \mathbf{v} \lesssim \mathbf{v}^T\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{mN}. \tag{24}$$

To show this result, we use the mesh properties (2) and (3), and the assumption that the diffusion coefficient and the penalty parameter are bounded above and below by positive constants (independent of h).

The main idea is to observe that A_Ω is an $N \times N$ block diagonal matrix with blocks of size $m \times m$, where the first row and column in every diagonal block is zero: this follows from the fact that $B_\Omega(\phi_k^i, \phi_\ell^{(j)}) = 0$ for $i \neq j$, and that the gradient of a piecewise constant basis function is zero. Altogether, A_Ω is of the following form:

$$A_\Omega = \left[\begin{array}{cc|c|cc|c|cc} 0 & 0 & & & & & & \\ 0 & A_\Omega^{(1)} & & & & & & \\ \hline & & \ddots & & & & & \\ \hline & & & 0 & 0 & & & \\ & & & 0 & A_\Omega^{(i)} & & & \\ \hline & & & & & \ddots & & \\ \hline & & & & & & 0 & 0 \\ & & & & & & 0 & A_\Omega^{(N)} \end{array}\right]. \tag{25}$$

As a consequence, we can treat the diagonal blocks individually using Lemma 3.2, and then combine the results (a similar strategy is used for the block diagonal D_σ).

To show (22), we also use the nature of $\pi_I = R^T R$, which is the projection operator onto the coarse space Range(R^T) that yields the best approximation in the 2-norm. As a result, the space Range($I - \pi_I$) is orthogonal to Range(R^T), where the latter corresponds to the piecewise constant basis functions. In particular, any $\mathbf{v} \in \text{Range}(I - \pi_I)$ is of the form:

$$\mathbf{v} = \left[\begin{array}{c} 0 \\ \hline \mathbf{v}_1 \\ \hline \vdots \\ \hline 0 \\ \hline \mathbf{v}_i \\ \hline \vdots \\ \hline 0 \\ \hline \mathbf{v}_N \end{array}\right]. \tag{26}$$

Using these ideas, we can now show Theorem 3.3:

Proof (of Theorem 3.3) Let $A_\Omega^{(i)}$ denote the result of deleting the first row and column in the i^{th} diagonal block in A_Ω, as indicated in (25). In other words:

$$\left(A_\Omega^{(i)}\right)_{\ell-1,k-1} = h^{2-d} B_\Omega^{(i)}(\phi_k^{(0)}, \phi_\ell^{(0)}),$$

for all $k, \ell = 2, ..., m$. Next, observe that $B_\Omega^{(0)}$ is independent of h and symmetric. Furthermore, for all higher-order polynomials $v \in \text{span}\{\phi_2^{(0)}, ..., \phi_m^{(0)}\} \setminus \{0\}$, the gradient of v is nonzero, which implies that $B_\Omega^{(0)}(v, v) > 0$. In other words, $B_\Omega^{(0)}$ is even positive-definite for the subspace under consideration. As a consequence, applying Lemma 3.2, we obtain a result similar to (22), but then for the individual diagonal blocks:

$$\mathbf{w}^T \mathbf{w} \lesssim \mathbf{w}^T A_\Omega^{(i)} \mathbf{w}, \qquad \forall \mathbf{w} \in \mathbb{R}^{m-1}, \qquad \forall i = 1, ..., N.$$

Using the notation in (26), this relations hold in particular for $\mathbf{w} = \mathbf{v}_i$, for all $i = 1, ..., N$. Summing over all i then yields:

$$\sum_{i=1}^N \mathbf{v}_i^T \mathbf{v}_i \lesssim \sum_{i=1}^N \mathbf{v}_i^T A_\Omega^{(i)} \mathbf{v}_i, \qquad \forall \mathbf{v} \in \text{Range}(I - \pi_I),$$

Using the notation in (26), this can be rewritten as (22), which then completes its proof. The relations (23) and (24) follow in a similar manner from Lemma 3.2 (without deleting the first row and column in each diagonal block). ■

4 Main result: scalability of both two-level methods

Using the auxiliary result obtained in the previous section, we can now show the main result of this paper: both two-level methods yield scalable convergence of the preconditioned CG method (independent of the mesh element diameter). This result is stated in Theorem 4.3 in Section 4.3. To show this result, we require two intermediate results, which are discussed in Section 4.1 and Section 4.2 respectively. Finally, Section 4.4 considers the special case of block Jacobi smoothing.

4.1 Intermediate result I: using the error iteration matrix

The first intermediate result bounds the condition number of the preconditioned system in terms of the eigenvalues of the so-called error iteration matrix. The following result holds for arbitrary SPD matrices A and P:

Lemma 4.1 *Suppose that A and P are SPD, and define $E = I - P^{-1}A$. Then, the condition number κ_2 (in the 2-norm) of $P^{-1}A$ can be bounded as follows (with $\lambda_{\min}(E) \leq \lambda_{\max}(E) < 1$):*

$$\kappa_2(P^{-1}A) \leq \frac{1 - \lambda_{\min}(E)}{1 - \lambda_{\max}(E)}. \tag{27}$$

To show this result, we make use of the following implication for positive constants c_1 and c_2 [22, p. 50][5]:

$$c_1 A \leq P \leq c_2 A \qquad \Rightarrow \qquad \kappa_2(P^{-1}A) \leq \frac{c_2}{c_1}. \tag{28}$$

PROOF (OF LEMMA 4.1) First, observe that E has the same spectrum as $A^{\frac{1}{2}} E A^{-\frac{1}{2}} = I - A^{\frac{1}{2}} P^{-1} A^{\frac{1}{2}}$, and that the latter is symmetric. Hence (cf. [16, p. 25]):

$$\lambda_{\max}(E) = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^T A^{\frac{1}{2}} E A^{-\frac{1}{2}} \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \overset{\mathbf{y}:=A^{-\frac{1}{2}}\mathbf{x}}{=} \max_{\mathbf{y}\neq 0} \frac{(A^{\frac{1}{2}}\mathbf{y})^T A^{\frac{1}{2}} E \mathbf{y}}{(A^{\frac{1}{2}}\mathbf{y})^T (A^{\frac{1}{2}}\mathbf{y})} = \max_{\mathbf{y}\neq 0} \frac{\mathbf{y}^T A E \mathbf{y}}{\mathbf{y}^T A \mathbf{y}}.$$

$$\lambda_{\min}(E) = ... = \min_{\mathbf{y}\neq 0} \frac{\mathbf{y}^T A E \mathbf{y}}{\mathbf{y}^T A \mathbf{y}}.$$

As a consequence, we may write:

$$\begin{aligned}
& \lambda_{\min}(E)\mathbf{y}^T A\mathbf{y} \leq \mathbf{y}^T A E \mathbf{y} \leq \lambda_{\max}(E)\mathbf{y}^T A\mathbf{y}, && \forall \mathbf{y}\neq 0 \\
\overset{E:=I-P^{-1}A}{\Rightarrow} \quad & \lambda_{\min}(E)\mathbf{y}^T A\mathbf{y} \leq \mathbf{y}^T (A - AP^{-1}A)\mathbf{y} \leq \lambda_{\max}(E)\mathbf{y}^T A\mathbf{y}, && \forall \mathbf{y}\neq 0 \\
\overset{\text{subtract } \mathbf{y}^T A\mathbf{y}}{\Rightarrow} \quad & (\lambda_{\min}(E)-1)\mathbf{y}^T A\mathbf{y} \leq -\mathbf{y}^T AP^{-1}A\mathbf{y} \leq (\lambda_{\max}(E)-1)\mathbf{y}^T A\mathbf{y}, && \forall \mathbf{y}\neq 0 \\
\overset{\text{times } -1}{\Rightarrow} \quad & (1-\lambda_{\min}(E))\mathbf{y}^T A\mathbf{y} \geq \mathbf{y}^T AP^{-1}A\mathbf{y} \geq (1-\lambda_{\max}(E))\mathbf{y}^T A\mathbf{y}, && \forall \mathbf{y}\neq 0 \\
\overset{\mathbf{w}:=A\mathbf{y}}{\Rightarrow} \quad & (1-\lambda_{\min}(E))\mathbf{w}^T A^{-1}\mathbf{w} \geq \mathbf{w}^T P^{-1}\mathbf{w} \geq (1-\lambda_{\max}(E))\mathbf{w}^T A^{-1}\mathbf{w}, && \forall \mathbf{w}\neq 0 \\
\overset{\text{[11, p. 398, 471]}}{\Rightarrow} \quad & \frac{1}{1-\lambda_{\min}(E)}\mathbf{v}^T A\mathbf{v} \leq \mathbf{v}^T P\mathbf{v} \leq \frac{1}{1-\lambda_{\max}(E)}\mathbf{v}^T A\mathbf{v}, && \forall \mathbf{v}\neq 0.
\end{aligned}$$

Implication (28) now yields (27), which completes the proof. ■

4.2 Intermediate result II: spectral bounds for the error iteration matrix

The second intermediate result concerns bounds for the eigenvalues of both two-level error iteration matrices. These bounds are obtained for *any* SPD matrix A, so the results below are not restricted to SIPG matrices. To specify the bounds, for any SPD matrix M, define

$$K_M := \sup_{\mathbf{v}\neq 0} \frac{\|(I-\pi_M)\mathbf{v}\|_M^2}{\|\mathbf{v}\|_A^2}.$$

Using this notation, the spectrum of the error iteration matrix $E_{\text{prec}} = I - P_{\text{prec}}^{-1}A$ can be bounded as follows (with $K_{\widetilde{M}_{\text{prec}}} \geq 1$):

$$0 \leq \lambda_{\min}(E_{\text{prec}}) \leq \lambda_{\max}(E_{\text{prec}}) \leq 1 - \frac{1}{K_{\widetilde{M}_{\text{prec}}}}. \tag{29}$$

This result for the two-level preconditioner follows as a special case from [7] (also cf. [22, p. 70-73]), and relies on the fact that M_{prec} is nonsingular, and that $M_{\text{prec}} + M_{\text{prec}}^T - A$ is SPD. In this section, we use a modified strategy to obtain similar spectral bounds for the deflation variant $E_{\text{defl}} = I - P_{\text{defl}}^{-1}A$:

Lemma 4.2 *The spectrum of the error iteration matrix E_{defl} above can be bounded as follows (with $K_{2M_{\text{defl}}} \geq 1$):*

$$-1 \leq \lambda_{\min}(E_{\text{defl}}) \leq \lambda_{\max}(E_{\text{defl}}) \leq 1 - \frac{1}{K_{2M_{\text{defl}}}}. \tag{30}$$

[5] Throughout this paper, for symmetrical matrices $M_1, M_2 \in \mathbb{R}^{n\times n}$, we write $M_1 \leq M_2$ to indicate that $\mathbf{v}^T M_1 \mathbf{v} \leq \mathbf{v}^T M_2 \mathbf{v}$ for all vectors $\mathbf{v} \in \mathbb{R}^n$; the notation $\geq$, $<$, and $>$ is used similarly.

To show this result, we use that M_{defl} and $2M_{\text{defl}} - A$ are SPD. Furthermore, we apply the following ideas: define $\bar{\pi}_A := A^{\frac{1}{2}} Q A^{\frac{1}{2}}$, and, for any SPD matrix M, set $\Theta_M := (I - \bar{\pi}_A)(I - A^{\frac{1}{2}} M^{-1} A^{\frac{1}{2}})(I - \bar{\pi}_A)$. Because $QAQ = Q$ [20, p. 1730], it can be verified that $\bar{\pi}_A = \bar{\pi}_A^2$ is a symmetric projection. Furthermore, E_{defl} has the same spectrum as $\Theta_{M_{\text{defl}}}$ [20, p. 1730], so we can analyze $\Theta_{M_{\text{defl}}}$ instead. At the same time, for any SPD matrix M:

$$M \geq A \qquad \Rightarrow \qquad \lambda_{\max}(\Theta_M) \leq 1 - \frac{1}{K_M}, \qquad K_M \geq 1. \tag{31}$$

This follows from [7] (also cf. [22, p. 71–73]). Altogether, Lemma 4.2 can now be shown as follows:

PROOF Because $\bar{\pi}_A$ is a symmetric projection, for any SPD matrices M_1 and M_2, it can be shown that

$$M_1 \leq M_2 \qquad \Rightarrow \qquad \Theta_{M_1} \leq \Theta_{M_2}. \tag{32}$$

This can be seen as follows:

$$\begin{aligned}
M_1 &\leq M_2, \\
M_1^{-1} &\overset{\text{[11, p. 398, 471]}}{\geq} M_2^{-1}, \\
A^{\frac{1}{2}} M_1^{-1} A^{\frac{1}{2}} &\geq A^{\frac{1}{2}} M_2^{-1} A^{\frac{1}{2}}, \\
I - A^{\frac{1}{2}} M_1^{-1} A^{\frac{1}{2}} &\leq I - A^{\frac{1}{2}} M_2^{-1} A^{\frac{1}{2}}, \\
\Theta_{M_1} &\leq \Theta_{M_2}.
\end{aligned}$$

Next, note that the assumption $2M_{\text{defl}} - A > 0$ implies that $M_{\text{defl}} \geq \frac{1}{2} A$. Using (32), we then obtain $\Theta_{M_{\text{defl}}} \geq \Theta_{\frac{1}{2}A} = -(I - \bar{\pi}_A)^2$. Since $\bar{\pi}_A$ is a projection, all eigenvalues of $(I - \bar{\pi}_A)^2 = I - \bar{\pi}_A$ are in $\{0, 1\}$. Hence, $\lambda_{\min}(\Theta_{M_{\text{defl}}}) \geq -1$. At the same time, it follows from (32) that $\Theta_{M_{\text{defl}}} \leq \Theta_{2M_{\text{defl}}}$. Applying (31) (using the assumption $2M_{\text{defl}} - A > 0$ once more) then implies that $\lambda_{\max}(\Theta_{M_{\text{defl}}}) \leq \lambda_{\max}(\Theta_{2M_{\text{defl}}}) \leq 1 - \frac{1}{K_{2M_{\text{defl}}}}$. Combining these two spectral bounds for $\Theta_{M_{\text{defl}}}$, and recalling that E_{defl} has the same spectrum as $\Theta_{M_{\text{defl}}}$, we arrive at (30), which then completes the proof. ■

4.3 Final main result: scalability of both two-level methods

Using the intermediate results in the previous sections, we can now show the main result of this paper: that both two-level methods yield scalable convergence of the CG method (independent of the mesh element diameter). This result has been shown by Dobrev et al. [6] for the preconditioning variant for $p = 1$. In this section, we use a similar strategy to extend these results for $p \geq 1$ and for the deflation variant.

Theorem 4.3 (Main result) *Both two-level methods yield scalable CG convergence in the sense that the condition number κ_2 (in the 2-norm) of the preconditioned system can be bounded independently of the maximum mesh element diameter h:*

$$\kappa_2(P_{\text{prec}}^{-1} A) \lesssim 1, \qquad \kappa_2(P_{\text{defl}}^{-1} A) \lesssim 1. \tag{33}$$

To show this result, we use coercivity (6), the mesh properties (2) and (3), and the assumption that the diffusion coefficient and the penalty parameter are bounded above and below by positive constants (independent of h). Regarding the smoothers, we use that M_{prec} is nonsingular, and that M_{defl}, $2M_{\text{defl}} - A$, and $M_{\text{prec}} + M_{\text{prec}}^T - A$ are SPD. Furthermore, we use the additional smoother requirements (14) and (15). It will be demonstrated in the next section that all of these smoother requirements are satisfied for (damped) block Jacobi smoothing.

The main idea is to combine Lemma 4.1, (29) and Lemma 4.2 to obtain:

$$\kappa_2(P_{\text{prec}}^{-1} A) \leq K_{\widetilde{M}_{\text{prec}}}, \qquad \kappa_2(P_{\text{defl}}^{-1} A) \leq 2 K_{2M_{\text{defl}}}. \tag{34}$$

The proof is then completed by showing that $K_{\widetilde{M}_{\text{prec}}}, K_{2M_{\text{defl}}} \lesssim 1$, for any smoother that satisfies the criteria above. This is established using the auxiliary result Theorem 3.3, and coercivity (6) in matrix form:

$$h^{d-2}\mathbf{v}^T \left(A_\Omega + A_\sigma\right)\mathbf{v} \lesssim \mathbf{v}^T A\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}. \tag{35}$$

Altogether, Theorem 4.3 can now be shown as follows.

PROOF (OF THEOREM 4.3) First, we will show that $K_{\widetilde{M}_{\text{prec}}} \lesssim 1$ (the proof that $K_{2M_{\text{defl}}} \lesssim 1$ is similar). For ease of notation, we will write $\widetilde{M}$ for $\widetilde{M}_{\text{prec}}$. The main idea is to show that $\|(I-\pi_{\widetilde{M}})\mathbf{v}\|_{\widetilde{M}} \lesssim \|\mathbf{v}\|_A$ for all $\mathbf{v}$: because $\pi_{\widetilde{M}}$ is a projection onto the coarse space Range(R^T) that yields the best approximation in the $\widetilde{M}$-norm, we can replace $\pi_{\widetilde{M}}$ by the suboptimal projection π_I:

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \leq \|(I-\pi_I)\mathbf{v}\|^2_{\widetilde{M}} = h^{d-2}\|(I-\pi_I)\mathbf{v}\|^2_{h^{2-d}\widetilde{M}}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Next, we apply the assumption on the smoother in (14):

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \lesssim h^{d-2}\|(I-\pi_I)\mathbf{v}\|_2^2, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Theorem 3.3 now gives (using (22)):

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \lesssim h^{d-2}\|(I-\pi_I)\mathbf{v}\|^2_{A_\Omega}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Because B_σ is SPSD, it follows that A_σ is SPSD. Hence:

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \lesssim h^{d-2}\|(I-\pi_I)\mathbf{v}\|^2_{A_\Omega+A_\sigma}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Next, we use coercivity (35):

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \lesssim \|(I-\pi_I)\mathbf{v}\|^2_A, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Finally, because π_I is a projection:

$$\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}} \lesssim \|\mathbf{v}\|^2_A, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Substitution of this relation into the definition of $K_{\widetilde{M}}$ yields:

$$K_{\widetilde{M}} = \sup_{\mathbf{v}\neq 0} \frac{\|(I-\pi_{\widetilde{M}})\mathbf{v}\|^2_{\widetilde{M}}}{\|\mathbf{v}\|^2_A} \lesssim 1,$$

A similar strategy, using (15) instead of (14), yields $K_{2M_{\text{defl}}} \lesssim 1$. Substitution of $K_{\widetilde{M}_{\text{prec}}}, K_{2M_{\text{defl}}} \lesssim 1$ into (34) now yields (33), which completes the proof of Theorem 4.3. ■

4.4 Special case: block Jacobi smoothing

This section demonstrates that Theorem 4.3 is valid for (damped) block Jacobi smoothing. To specify this result, suppose that M_{BJ} is the block Jacobi smoother with blocks of size $m \times m$. Next, consider the specific choices $M_{\text{prec}}, M_{\text{defl}} = \omega^{-1}M_{\text{BJ}}$ with damping parameter $\omega > 0$ (independent of h). We assume that $\omega \leq 1$, with $\omega < 1$ strictly for the preconditioning variant. Furthermore, we assume that the mesh can be colored by two colors, i.e. that the mesh can be represented by a graph whose vertices can be colored such that connected vertices do not have the same color. This is the case e.g. for structured rectangular meshes. Alternatively, we could assume that the damping parameter ω is sufficiently small. We do not discuss this latter alternative strategy further. Altogether, all smoother requirements for Theorem 4.3 are satisfied for (damped) block Jacobi smoothing:

Corollary 4.4 *Suppose that the mesh can be colored by two colors. Then, Theorem 4.3 applies for the damped block Jacobi smoothers M_{prec} and M_{defl} above, i.e. both two-level methods yield scalable CG convergence in the sense that the condition number κ_2 (in the 2-norm) of the preconditioned system can be bounded independently of the maximum mesh element diameter h:*

$$\kappa_2(P^{-1}_{\text{prec}}A) \lesssim 1, \qquad \kappa_2(P^{-1}_{\text{defl}}A) \lesssim 1.$$

This result follows immediately from Theorem 4.3 once we have verified that the conditions (10), (13), (14), and (15) are satisfied for the (SPD) damped block Jacobi smoothers under consideration. In other words, writing $M := \omega^{-1} M_{\mathrm{BJ}}$, we need to show:

$$2M - A > 0, \tag{36}$$

$$h^{2-d} \mathbf{v}^T M \mathbf{v} \lesssim \mathbf{v}^T \mathbf{v}, \qquad \forall \mathbf{v} \in \mathrm{Range}(I - \pi_I), \tag{37}$$

$$h^{2-d} \mathbf{v}^T \widetilde{M} \mathbf{v} \lesssim \mathbf{v}^T \mathbf{v}, \qquad \forall \mathbf{v} \in \mathrm{Range}(I - \pi_I), \tag{38}$$

for all $\omega \leq 1$, with $\omega < 1$ strictly for (38). We treat each relation separately.

To show (36), the main idea is to observe that $2M_{\mathrm{BJ}} - A$ is the same as A, except that the off-diagonal blocks have an additional minus sign. Using this property, the proof of (36) is completed using that the mesh can be colored by two colors:

Proof (of (36)) Choose arbitrary $\mathbf{v} \in \mathbb{R}^{Nm}$, and write

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_N \end{pmatrix}, \qquad \text{with } \mathbf{v}_1, ..., \mathbf{v}_N \in \mathbb{R}^m.$$

Without loss of generality, we may assume that $\omega = 1$. In that case, $2M - A$ is the same as A, except that the off-diagonal blocks have an additional minus sign. Hence, we may write:

$$\mathbf{v}^T (2M - A) \mathbf{v} = \sum_{i=1}^{N} \mathbf{v}_i^T A_{ii} \mathbf{v}_i - \sum_{i=1}^{N} \sum_{j \neq i} \mathbf{v}_i^T A_{ij} \mathbf{v}_j.$$

Next, suppose that $W \subset \{1, 2, ..., N\}$ is the collection of indices corresponding to one of the two colors with which the mesh can be colored. Define the complement $W^C = \{1, 2, ..., N\} \setminus W$, i.e. the collection of indices corresponding to the second color. Because $W \cup W^C = \{1, 2, ..., N\}$, we may write:

$$\mathbf{v}^T (2M - A) \mathbf{v} = \sum_{i=1}^{N} \mathbf{v}_i^T A_{ii} \mathbf{v}_i - \sum_{i \in W} \sum_{j \neq i} \mathbf{v}_i^T A_{ij} \mathbf{v}_j - \sum_{i \in W^C} \sum_{j \neq i} \mathbf{v}_i^T A_{ij} \mathbf{v}_j.$$

Because $A_{ij} = 0$ if $i \neq j$ are either both in W or both in W^C, it follows that:

$$\mathbf{v}^T (2M - A) \mathbf{v} = \sum_{i=1}^{N} \mathbf{v}_i^T A_{ii} \mathbf{v}_i - \sum_{i \in W} \sum_{j \in W^C \setminus \{i\}} \mathbf{v}_i^T A_{ij} \mathbf{v}_j - \sum_{i \in W^C} \sum_{j \in W \setminus \{i\}} \mathbf{v}_i^T A_{ij} \mathbf{v}_j.$$

Next, define:

$$\mathbf{w} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N \end{pmatrix} \in \mathbb{R}^{Nm}, \qquad \mathbf{w}_j = \begin{cases} -\mathbf{v}_j, & \text{for } j \in W, \\ \mathbf{v}_j, & \text{for } j \in W^C. \end{cases}$$

Substitution yields:

$$\mathbf{v}^T (2M - A) \mathbf{v} = \sum_{i=1}^{N} \mathbf{w}_i^T A_{ii} \mathbf{w}_i + \sum_{i \in W} \sum_{j \in W^C \setminus \{i\}} \mathbf{w}_i^T A_{ij} \mathbf{w}_j + \sum_{i \in W^C} \sum_{j \in W \setminus \{i\}} \mathbf{w}_i^T A_{ij} \mathbf{w}_j.$$

As before, we can use that $W \cup W^C = \{1, 2, ..., N\}$, and that $A_{ij} = 0$ if $i \neq j$ are either both in W or both in W^C:

$$\mathbf{v}^T (2M - A) \mathbf{v} = \sum_{i=1}^{N} \mathbf{w}_i^T A_{ii} \mathbf{w}_i + \sum_{i=1}^{N} \sum_{j \neq i} \mathbf{w}_i^T A_{ij} \mathbf{w}_j = \mathbf{w}^T A \mathbf{w}.$$

And because A is SPD, it now follows that $2M - A$ is SPD. This completes the proof of (36). ■

To show (37), the main idea is to use Theorem 3.3 and the following property (cf. [6, p. 760] and [12, p. 4]):

$$0 < B(v, v) \lesssim B_\Omega(v, v) + B_\sigma(v, v), \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}. \tag{39}$$

PROOF (OF (37)) Without loss of generality, we may assume that $\omega = 1$. Next, recall the notation introduced in the beginning of Section 3.3. Additionally, similar to D_σ, let D_r be the result of extracting the diagonal blocks of size $m \times m$ from A_r. Using this notation, and the fact that A_Ω is a block diagonal matrix with blocks of size $m \times m$, we may write:

$$h^{2-d}M = A_\Omega + D_\sigma + D_r.$$

Next, we write (39) in matrix form:

$$0 < \mathbf{v}^T A \mathbf{v} \lesssim \mathbf{v}^T \left(h^{d-2} A_\Omega + h^{d-2} A_\sigma \right) \mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Because this relation is also true when considering the diagonal blocks only, we may write:

$$h^{2-d}\mathbf{v}^T M \mathbf{v} \lesssim \mathbf{v}(A_\Omega + D_\sigma)\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{Nm}.$$

Application of Theorem 3.3 now yields (37), which completes the proof. ■

To show (38), we combine the previous results (36) and (37):

PROOF (OF (38)) Using (36), and the fact that $\omega < 1$ strictly, it can be shown that

$$\widetilde{M} \leq \frac{1}{2(1-\omega)} M. \tag{40}$$

This can be seen as follows:

$$\begin{aligned}
2\omega M - A &\geq 0, \\
2M - A = (2-2\omega)M + \underbrace{2\omega M - A}_{\geq 0} &\geq (2-2\omega)M, \\
(2M-A)^{-1} &\overset{\text{[11, p. 398, 471]}}{\leq} \frac{1}{2(1-\omega)} M^{-1}, \\
\underbrace{M(2M-A)^{-1}M}_{=:\widetilde{M}} &\leq \frac{1}{2(1-\omega)} M.
\end{aligned}$$

Combining this relation with (37), it now follows that

$$h^{2-d}\mathbf{v}^T \widetilde{M} \mathbf{v} \overset{(40)}{\leq} \frac{1}{2(1-\omega)} h^{2-d}\mathbf{v}^T M \mathbf{v} \overset{(37)}{\lesssim} \mathbf{v}^T \mathbf{v}, \qquad \forall \mathbf{v} \in \text{Range}(I - \pi_I).$$

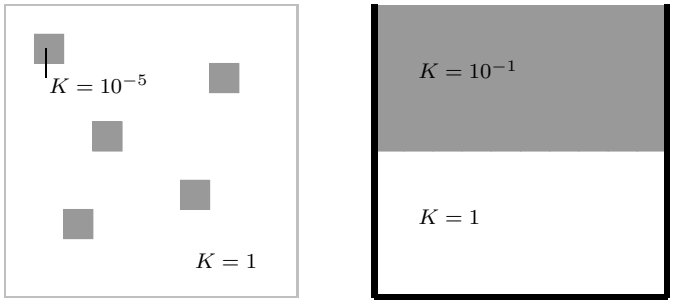
This completes the proof of (38). ■

5 Numerical results

The previous section demonstrated theoretically that both two-level methods yield scalable convergence of the CG method. In this section, we extend the numerical support in [19] for this result by studying two test problems with strong variations in the coefficients.

Test cases We consider two diffusion problems of the form (1) on the domain $[0,1]^2$, as illustrated in Figure 1 (if we subdivide the domain into 10×10 equally sized squares, the diffusion coefficient is constant within each square). The first problem is a bubbly flow problem with large jumps in the coefficients, inspired by [20]. The second problem involves homogeneous Neumann boundary conditions (indicated by the black lines in Figure 1). For both problems, the Dirichlet boundary conditions and the source term f are chosen such that the exact solution reads $u(x,y) = \cos(10\pi x)\cos(10\pi y)$. We stress that this choice does not impact the matrix or the performance of the linear solver, as we use random start vectors (see below).

Experimental setup All model problems are discretized by means of the SIPG method as discussed in Section 2.1. The penalty parameter is chosen diffusion-dependent, $\sigma = 20K$ (using the largest trace value of K at the discontinuities), as motivated by [19]. The difference with a constant penalty parameter is discussed at the end of this section. Furthermore, we use a uniform Cartesian mesh with $N = n \times n$ elements with $n = 40, 80, 160, 320$, and monomial basis functions



Bubbly flow Neumann BCs

Figure 1: Illustration of the test cases

with polynomial degree $p = 2, 3$ (results for $p = 1$ are similar though). As a result, the largest problems have over 10^6 degrees of freedom.

The resulting linear systems are solved by means of the preconditioned CG method, combined with either the two-level preconditioner or the corresponding ADEF2 deflation variant, as discussed in Section 2.2. Furthermore, we use (damped) block Jacobi smoothing (cf. Section 4.4). For the damping parameter we consider $\omega = 1$ for the deflation variant, and both $\omega = 1$ and $\omega = 0.7$ for the preconditioner, as motivated by [19]. Diagonal scaling is applied as a pre-processing step in all cases, and the same random start vector $\mathbf{x}_0$ is used for all problems of the same size. For the stopping criterion we use: $\frac{\|r_k\|_2}{\|b\|_2} \leq 10^{-6}$, where r_k is the residual after the k^{th} iteration. Coarse systems, involving the SIPG matrix A_0 with polynomial degree $p = 0$, are solved directly. However, a more efficient strategy has been studied in [19]. In any case, the coarse matrix A_0 is quite similar to a central difference matrix, for which very efficient solvers are readily available.

Results Table 1 and Table 3 display the results in terms of the number of CG iterations required for convergence. The corresponding computational times are provided in Table 2 and Table 4. It can be seen that both two-level methods yield fast and scalable convergence. Without damping, deflation is the most efficient. When a suitable damping value is known, the preconditioning variant performs comparable to deflation.

degree	p=2				p=3			
mesh	N=40^2	N=80^2	N=160^2	N=320^2	N=40^2	N=80^2	N=160^2	N=320^2
Prec., 2x BJ	41	42	43	44	55	56	57	58
Prec., 2x BJ ($\omega = 0.7$)	31	31	32	32	33	34	35	35
Defl., 1x BJ	41	39	40	41	45	45	45	46

Table 1: Bubbly flow: # CG Iterations

degree	p=2				p=3			
mesh	N=40^2	N=80^2	N=160^2	N=320^2	N=40^2	N=80^2	N=160^2	N=320^2
Prec., 2x BJ	0.09	0.60	3.07	15.42	0.38	1.75	7.95	36.74
Prec., 2x BJ ($\omega = 0.7$)	0.07	0.45	2.30	11.35	0.23	1.07	4.89	22.69
Defl., 1x BJ	0.07	0.45	2.38	12.47	0.21	1.01	4.77	22.39

Table 2: Bubbly flow: CPU time in seconds

Influence of the penalty parameter The results in this section have been established using a diffusion-dependent penalty parameter $\sigma = 20K$. A common alternative value is to choose the penalty parameter constant, e.g. $\sigma = 20$. Both options have been compared numerically in [19] for a diffusion problem with five layers where either $K = 1$ or $K = 10^{-3}$ in each layer. In

degree	p=2				p=3			
mesh	$N=40^2$	$N=80^2$	$N=160^2$	$N=320^2$	$N=40^2$	$N=80^2$	$N=160^2$	$N=320^2$
Prec., 2x BJ	45	45	45	45	59	59	60	60
Prec., 2x BJ ($\omega = 0.7$)	34	35	36	36	36	37	37	38
Defl., 1x BJ	47	47	47	47	49	49	50	50

Table 3: Neumann BCs: # CG Iterations

degree	p=2				p=3			
mesh	$N=40^2$	$N=80^2$	$N=160^2$	$N=320^2$	$N=40^2$	$N=80^2$	$N=160^2$	$N=320^2$
Prec., 2x BJ	0.10	0.65	3.31	16.01	0.40	1.86	8.53	37.77
Prec., 2x BJ ($\omega = 0.7$)	0.08	0.51	2.67	12.97	0.24	1.18	5.33	24.33
Defl., 1x BJ	0.07	0.53	2.82	14.37	0.23	1.09	5.23	24.52

Table 4: Neumann BCs: CPU time in seconds

that reference, it has been found that both two-level methods perform significantly worse for the constant penalty parameter: up to 150 times more iterations were required for the preconditioning variant, and up to 18 times more iterations for the deflation variant. Furthermore, the convergence did not seem scalable for the meshes under consideration (i.e. more iterations for finer meshes for $\sigma = 20$).

This does *not* contradict the theoretical scalability result in this paper: this result implies that the number of iterations can be bounded independently of the mesh element diameter, *not* that the number of iterations is the same for each mesh. In other words, for sufficiently fine meshes, the scalable convergence should become apparent for the problem above. This hypothesis has been verified by studying a similar problem with smaller jumps in the coefficients (either $K = 1$ or $K = 0.5$ in each of the five layers). For this 'reduced' problem, we observe only a slight increase in the number of iterations, and the magnitude of this increase reduces with the mesh element diameter.

Altogether, scalable convergence is obtained for both a constant and a diffusion-dependent penalty parameter. However, a diffusion-dependent penalty parameter yields significantly faster results for problems with strong variations in the coefficients.

6 Conclusion

This paper is focused on a two-level preconditioner proposed in [6] and the corresponding BNN (ADEF2) deflation variant for linear SIPG systems. For both two-level methods, we have found that the condition number of the preconditioned system can be bounded independently of the mesh element diameter, implying scalable CG convergence. This result is valid for polynomial degree $p \geq 1$. We have verified that the restrictions on the smoother are satisfied for block Jacobi smoothing. Numerical experiments with strong variations in the coefficients further support our main result. Future research could focus on a theoretical *comparison* of both two-level methods and on more advanced, larger scale numerical test cases.

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(1):21–54, 2007.

[2] P. Bastian, M. Blatt, and R. Scheichl. Algebraic multigrid for discontinuous Galerkin discretizations of heterogeneous elliptic problems. *Numer. Linear Algebra Appl.*, 19(2):367–388, 2012.

[3] S. C. Brenner and J. Zhao. Convergence of multigrid algorithms for interior penalty methods. *Appl. Numer. Anal. Comput. Math.*, 2(1):3–18, 2005.

[4] P. Castillo. Performance of discontinuous Galerkin methods for elliptic PDEs. *SIAM J. Sci. Comput.*, 24(2):524–547, 2002.

[5] P. G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[6] 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(9):753–770, 2006.

[7] R. D. Falgout, P. S. Vassilevski, and L. T. Zikatanov. On two-grid convergence estimates. *Numer. Linear Algebra Appl.*, 12(5-6):471–494, 2005.

[8] 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(4):1343–1365 (electronic), 2001.

[9] 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(1):92–113, 2005.

[10] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95(3):527–550, 2003.

[11] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1988.

[12] K. Johannsen. A symmetric smoother for the nonsymmetric interior penalty discontinuous galerkin discretization. Technical Report ICES Report 05-23, University of Texas at Austin, 2005.

[13] P.-O. Persson and J. Peraire. Newton-GMRES preconditioning for discontinuous Galerkin discretizations of the Navier-Stokes equations. *SIAM J. Sci. Comput.*, 30(6):2709–2733, 2008.

[14] F. Prill, M. Lukáčová-Medviďová, and R. Hartmann. Smoothed aggregation multigrid for the discontinuous Galerkin method. *SIAM J. Sci. Comput.*, 31(5):3503–3528, 2009.

[15] B. Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations*, volume 35 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. Theory and implementation.

[16] Y. Saad. Iterative methods for sparse linear systems. This is a revised version of the book published in 1996 by PWS Publishing, Boston. It can be downloaded from http://www-users.cs.umn.edu/ saad/books.html, 2000.

[17] Y. Saad and B. Suchomel. ARMS: an algebraic recursive multilevel solver for general sparse linear systems. *Numer. Linear Algebra Appl.*, 9(5):359–378, 2002.

[18] S. J. Sherwin, R. M. Kirby, J. Peiró, R. L. Taylor, and O. C. Zienkiewicz. On 2D elliptic discontinuous Galerkin methods. *Internat. J. Numer. Methods Engrg.*, 65(5):752–784, 2006.

[19] P. van Slingerland and C. Vuik. Fast linear solver for pressure computation in layered domains. Technical Report 12-10, Delft University of Technology, 2012.

[20] J. M. Tang, S. P. MacLachlan, R. Nabben, and C. Vuik. A comparison of two-level preconditioners based on multigrid and deflation. *SIAM J. Matrix Anal. Appl.*, 31(4):1715–1739, 2009/10.

[21] J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga. Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods. *J. Sci. Comput.*, 39(3):340–370, 2009.

[22] P. S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008. Matrix-based analysis and algorithms for solving finite element equations.

[23] Jinchao Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34(4):581–613, 1992.