Numer. Math. (2014) 128:193–216
DOI 10.1007/s00211-014-0607-3

Numerische
Mathematik

Multigrid methods for saddle point problems: Stokes and Lamé systems

Susanne C. Brenner · Hengguang Li · Li-Yeng Sung

Received: 21 April 2013 / Revised: 3 November 2013 / Published online: 31 January 2014

Abstract We develop new multigrid methods for a class of saddle point problems that include the Stokes system in fluid flow and the Lamé system in linear elasticity as special cases. The new smoothers in the multigrid methods involve optimal preconditioners for the discrete Laplace operator. We prove uniform convergence of the W-cycle algorithm in the energy norm and present numerical results for W-cycle and V-cycle algorithms.

Mathematics Subject Classfication (1991) Primary 65N55 · 65F10 · 65N30; Secondary 76D07 · 74B05

1 Introduction

In this paper we consider multigrid methods for a class of saddle point problems that include the stationary Stokes system in fluid flow with the no-slip boundary condition

The work of the first and third authors was supported in part by the National Science Foundation under Grant Nos. DMS-10-16332 and DMS-13-19172. The work of the second author was supported in part by the National Science Foundation under Grant No. DMS-11-58839. This work was also supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation.

S. C. Brenner (✉) · L.-Y. Sung
Department of Mathematics and Center for Computation and Technology,
Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: brenner@math.lsu.edu

L.-Y. Sung
e-mail: sung@math.lsu.edu

H. Li
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA
e-mail: hli@math.wayne.edu

and the Lamé system in linear elasticity with the homogeneous displacement boundary condition as special cases. We will follow standard notation for differential operators and Sobolev spaces that are found for example in [16,19], and we will use boldfaced letters to denote vector functions. Throughout the paper Ω is a bounded polyhedral domain in $\mathbb{R}^d$ $(d = 2, 3)$.

The Stokes problem (cf. [18,24]) is to find $(\boldsymbol{u}, p) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$ such that

$$\nu \int_\Omega \nabla \boldsymbol{u} : \nabla \boldsymbol{v}\, dx - \int_\Omega (\nabla \cdot \boldsymbol{v}) p\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx \qquad \forall\, \boldsymbol{v} \in [H_0^1(\Omega)]^d, \tag{1.1a}$$

$$- \int_\Omega (\nabla \cdot \boldsymbol{u}) q\, dx = 0 \qquad \forall\, q \in L_2^0(\Omega), \tag{1.1b}$$

where $\boldsymbol{u}$ is the fluid velocity, p is the pressure, ν is the kinematic viscosity, $\boldsymbol{f} \in [L_2(\Omega)]^d$ is the density of body force, $L_2^0(\Omega) = \{q \in L_2(\Omega) : \int_\Omega q\, dx = 0\}$, and the colon stands for the Frobenius product between matrices.

The Lamé problem (cf. [18,21]) is to find $(\boldsymbol{u}, p) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$ such that

$$2\mu \int_\Omega \boldsymbol{\epsilon}(\boldsymbol{u}) : \boldsymbol{\epsilon}(\boldsymbol{v})\, dx - \int_\Omega (\nabla \cdot \boldsymbol{v}) p\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx \qquad \forall\, \boldsymbol{v} \in [H_0^1(\Omega)]^d, \tag{1.2a}$$

$$- \int_\Omega (\nabla \cdot \boldsymbol{u}) q\, dx - \frac{1}{\lambda} \int_\Omega pq\, dx = 0 \qquad \forall\, q \in L_2^0(\Omega), \tag{1.2b}$$

where $\boldsymbol{u}$ is the displacement, $p = -\lambda(\nabla \cdot \boldsymbol{u})$, $\boldsymbol{f} \in [L_2(\Omega)]^d$ is the load density, the strain tensor $\boldsymbol{\epsilon}(\boldsymbol{v}) = \frac{1}{2}[(\nabla \boldsymbol{v}) + (\nabla \boldsymbol{v})^t]$ is the symmetrized gradient of $\boldsymbol{v}$, and μ and λ are the Lamé constants.

These two problems are special cases of the following saddle point problem: Find $(\boldsymbol{u}, p) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$ such that

$$\mathcal{B}\big((\boldsymbol{u}, p), (\boldsymbol{v}, q)\big) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \forall\, (\boldsymbol{v}, q) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega), \tag{1.3}$$

where

$$\mathcal{B}\big((\boldsymbol{w}, r), (\boldsymbol{v}, q)\big) = a(\boldsymbol{w}, \boldsymbol{v}) + b(\boldsymbol{v}, r) + b(\boldsymbol{w}, q) - t(r, q). \tag{1.4}$$

Here $(\cdot, \cdot)$ is the L_2 inner product, t is a nonnegative number, the bilinear form $b(\cdot, \cdot)$ on $[H_0^1(\Omega)]^d \times L_2^0(\Omega)$ is given by

$$b(\boldsymbol{v}, q) = - \int_\Omega (\nabla \cdot \boldsymbol{v}) q\, dx, \tag{1.5}$$

and $a(\cdot, \cdot)$ is a symmetric bounded and coercive bilinear form on $[H_0^1(\Omega)]^d$, i.e., there exist positive constants γ_b and γ_c such that

$$a(\boldsymbol{v}, \boldsymbol{w}) \leq \gamma_b \|\boldsymbol{v}\|_{H^1(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)} \qquad \forall\, \boldsymbol{v}, \boldsymbol{w} \in [H_0^1(\Omega)]^d, \tag{1.6}$$

$$a(\boldsymbol{v}, \boldsymbol{v}) \geq \gamma_c \|\boldsymbol{v}\|^2_{H^1(\Omega)} \qquad \forall\, \boldsymbol{v} \in [H_0^1(\Omega)]^d. \tag{1.7}$$

The Stokes problem (1.1) corresponds to the choice of $t = 0$ and

$$a(\boldsymbol{w}, \boldsymbol{v}) = \nu \int_\Omega \nabla \boldsymbol{w} : \nabla \boldsymbol{v}\, dx,$$

while the Lamé problem (1.2) corresponds to the choice of $t = 1/\lambda$ and

$$a(\boldsymbol{w}, \boldsymbol{v}) = 2\mu \int_\Omega \boldsymbol{\epsilon}(\boldsymbol{w}) : \boldsymbol{\epsilon}(\boldsymbol{v})\, dx.$$

The coercivity of the bilinear form $a(\cdot, \cdot)$ follows from a Poincaré–Friedrichs inequality (Stokes problem) and a Korn inequality (Lamé problem).

Remark 1.1 The problem (1.3) can be posed for $\boldsymbol{f} \in [H^{-1}(\Omega)]^d$, where the integral on the right-hand side is replaced by the duality pairing $\langle \boldsymbol{f}, \boldsymbol{v} \rangle$.

The problem (1.3) can be discretized by a stable finite element method based on the finite element spaces V_h ($\subset [H_0^1(\Omega)]^d$) and Q_h ($\subset L_2^0(\Omega)$) that satisfy the inf-sup condition. Our goal is to develop multigrid methods for the resulting finite dimensional saddle point problem that are uniformly convergent in the energy norm $\|\cdot\|$ defined by

$$\|(\boldsymbol{v}, q)\| = \|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)},$$

i.e., the multigrid method is a contraction with respect to the energy norm and the contraction number is bounded away from 1 on all grid levels. In fact we will show that the contraction number for the Stokes (resp. Lamé) problem satisfies an estimate of the form C/m^{α_S} (resp. C/m^{α_L}), where m is the number of smoothing steps, C is a positive constant independent of grid levels, and $\alpha_S \in (0, 1]$ (resp. $\alpha_L \in (0, 1]$) is the index of elliptic regularity for the Stokes (resp. Lamé) problem that can be taken to be 1 if Ω is convex.

Multigrid methods for the saddle point problem (1.3) have been investigated in [6,11–13,34,36,42–44,47]. However the multigrid convergence results in these papers are established with respect to norms different from the energy norm. The analyses in these papers also require Ω to be convex and the contraction number bounds in some of these papers are of the form $C/\sqrt{m}$.

The key ingredients of our multigrid methods are two new smoothers. Let $\mathbb{B}_h : V_h \times Q_h \longrightarrow V_h \times Q_h$ be the linear operator that represents $\mathcal{B}(\cdot, \cdot)$ with respect to an inner product $[\cdot, \cdot]_h$ that satisfies

$$[(\boldsymbol{v}_h, q_h), (\boldsymbol{v}_h, q_h)]_h \approx \|\boldsymbol{v}_h\|^2_{L_2(\Omega)} + h^2 \|q_h\|^2_{L_2(\Omega)} \qquad \forall\, (\boldsymbol{v}_h, q_h) \in V_h \times Q_h. \tag{1.8}$$

In the post-smoothing step, we take the smoother for an equation of the form

$$\mathbb{B}_h(\boldsymbol{w}_h, r_h) = (\boldsymbol{g}_h, z_h)$$

to be the Richardson relaxation for the equivalent equation

$$\mathbb{B}_h\mathbb{S}_h\mathbb{B}_h(\boldsymbol{w}_h, r_h) = \mathbb{B}_h\mathbb{S}_h(\boldsymbol{g}_h, z_h),$$

where the operator $\mathbb{S}_h$ is symmetric positive definite (SPD) and

$$[\mathbb{B}_h\mathbb{S}_h\mathbb{B}_h(\boldsymbol{v}_h, q_h), (\boldsymbol{v}_h, q_h)]_h \approx \|\boldsymbol{v}_h\|^2_{H^1(\Omega)} + \|q_h\|^2_{L_2(\Omega)}. \tag{1.9}$$

The construction of $\mathbb{S}_h$ involves an optimal preconditioner for the discrete Laplace operator and the property (1.9) relies on the inf-sup condition for V_h and Q_h.

In the pre-smoothing step, we use a smoother that is dual (with respect to the bilinear form $\mathcal{B}(\cdot,\cdot)$) to the post-smoother, which allows us to deduce immediately its smoothing property from the smoothing property of the post-smoother.

Remark 1.2 Because of (1.8) and (1.9), we can use the operator $\mathbb{B}_h\mathbb{S}_h\mathbb{B}_h$ to define mesh-dependent norms (cf. (4.1)) that are related to Sobolev norms and obtain smoothing and approximation properties (cf. Lemmas 5.1, 5.4) that are similar to those of second order scalar elliptic problems. Therefore we can apply multigrid techniques originally invented for SPD problems [7,10,28,33,40] to handle saddle point problems on general polygonal domains where the full elliptic regularity is not available.

Remark 1.3 Since an optimal preconditioner for the discrete Laplace operator is used in the smoothing steps, our multigrid method can also be viewed as a preconditioned iterative method for saddle point problems. Other preconditioned iterative methods for saddle problems are discussed in [4,22,23,31,41] and the references therein. (See Sect. 7 for further discussions.)

The rest of the paper is organized as follows. In Sect. 2 we consider various aspects of the saddle point problem (1.3) that are relevant to the multigrid convergence analysis. The new multigrid methods are introduced in Sect. 3. In Sect. 4 we establish properties of certain mesh-dependent norms that are useful for the convergence analysis of the W-cycle algorithm in Sect. 5. Numerical results that illustrate the performance of the multigrid methods on two dimensional domains are presented in Sect. 6 and we end with some concluding remarks in Sect. 7. Appendix contains the proof of an elliptic regularity estimate for the two dimensional Lamé system that is robust with respect to the Lamé constant λ.

2 The saddle point problem (1.3)

In this section we discuss conditions on the saddle point problem (1.3) under which the convergence of the W-cycle multigrid method will be established. We begin with properties of the bilinear forms. Then we will consider elliptic regularity and finite element discretizations.

2.1 Properties of the bilinear forms

From (1.4), (1.5) and (1.6) we immediately see that the bilinear form $\mathcal{B}(\cdot,\cdot)$ is bounded with respect to the energy norm:

$$\mathcal{B}\big((\boldsymbol{v},q),(\boldsymbol{w},r)\big) \le C_{\gamma_b,t}(\|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)})(\|\boldsymbol{w}\|_{H^1(\Omega)} + \|r\|_{L_2(\Omega)}) \tag{2.1}$$

for all $(\boldsymbol{v},q),(\boldsymbol{w},r) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$. We will assume that

$$0 \le t \le t_* < \infty, \tag{2.2}$$

which is the relevant range for the Stokes and Lamé problems.

The bilinear form $b(\cdot,\cdot)$ satisfies the following inf-sup condition (cf. [24]):

$$\inf_{0\neq q\in L_2^0(\Omega)} \sup_{\mathbf{0}\neq \boldsymbol{v}\in[H_0^1(\Omega)]^d} \frac{b(\boldsymbol{v},q)}{|\boldsymbol{v}|_{H^1(\Omega)}\|q\|_{L_2(\Omega)}} \ge \beta_c > 0, \tag{2.3}$$

which together with (1.6) and (1.7) implies that (1.3) is well-posed (cf. [18, Sect. II.1.2]).

2.2 Elliptic regularity

The theory of elliptic regularity for the Stokes and Lamé problems can be found for example in the monographs [20,25–27,29,32].

For the Stokes problem (1.1), there exists $\alpha_S \in \left(\frac{1}{2},1\right]$ such that the solution $(\boldsymbol{u},p)$ belongs to $[H^{1+\alpha}(\Omega)]^d \times H^\alpha(\Omega)$ for $0 \le \alpha \le \alpha_S$ if $\boldsymbol{f} \in [H^{-1+\alpha}(\Omega)]^d$, and we have

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + \|p\|_{H^\alpha(\Omega)} \le C_{\Omega,\nu_0}\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} \tag{2.4}$$

that is valid for $\nu \ge \nu_0 > 0$. The index of elliptic regularity $\alpha_S = 1$ if Ω is convex and $\alpha_S < 1$ if Ω is nonconvex.

For the Lamé problem (1.2), there exists $\alpha_L \in \left(\frac{1}{2},\alpha_S\right]$ such that the solution $(\boldsymbol{u},p)$ belongs to $[H^{1+\alpha}(\Omega)]^d \times H^\alpha(\Omega)$ for $0 \le \alpha \le \alpha_L$ if $\boldsymbol{f} \in [H^{-1+\alpha}(\Omega)]^d$, and we have

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + \|p\|_{H^\alpha(\Omega)} \le C_{\Omega,\mu_0,\mu_1,\lambda_0,\lambda_1}\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} \tag{2.5}$$

that is valid for $0 < \mu_0 \le \mu \le \mu_1 < \infty$ and $0 < \lambda_0 \le \lambda \le \lambda_1 < \infty$. The index of elliptic regularity $\alpha_L = 1$ if Ω is convex and $\alpha_L \le \alpha_S < 1$ if Ω is nonconvex.

In the two dimensional case, we can take advantage of the results in [1] to derive a robust regularity estimate with respect to λ. The proof of the following proposition is given in the Appendix.

Proposition 2.1 *Let μ_0, μ_1 and λ_0 be positive numbers. There exists a positive constant $C_{\Omega,\mu_0,\mu_1,\lambda_0}$ such that the estimate*

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + \|p\|_{H^\alpha(\Omega)} \le C_{\Omega,\mu_0,\mu_1,\lambda_0}\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} \tag{2.6}$$

holds for the solution $(\boldsymbol{u}, p)$ *of* (1.2) *with* $d = 2$, *provided* $\mu_0 \le \mu \le \mu_1$, $\lambda_0 \le \lambda < \infty$ *and* $0 \le \alpha \le \alpha_L$.

In view of (2.4) and (2.5), we will assume from here on that the solution of (1.3) satisfies

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + \|p\|_{H^{\alpha}(\Omega)} \le C_E \|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} \tag{2.7}$$

for some $\alpha \in \left(\frac{1}{2}, 1\right]$, which is crucial for establishing the approximation property of multigrid methods.

2.3 Finite element methods

Let $\mathcal{T}_h$ be a triangulation of Ω, where the mesh parameter h is proportional to the maximum of the diameters of the element domains in the triangulation. Let $V_h \subset [H_0^1(\Omega)]^d$ and $Q_h \subset L_2^0(\Omega)$ be finite element spaces associated with $\mathcal{T}_h$. The discrete problem for (1.3) is to find $(\boldsymbol{u}_h, p_h) \in V_h \times Q_h$ such that

$$\mathcal{B}\big((\boldsymbol{u}_h, p_h), (\boldsymbol{v}, q)\big) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \forall\, (\boldsymbol{v}, q) \in V_h \times Q_h. \tag{2.8}$$

We assume the finite element pair $V_h \times Q_h$ satisfies a discrete inf-sup condition:

$$\inf_{0 \ne q \in Q_h} \sup_{\boldsymbol{0} \ne \boldsymbol{v} \in V_h} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_{H^1(\Omega)} \|q\|_{L_2(\Omega)}} \ge \beta_d > 0, \tag{2.9}$$

where β_d is independent of h.

It follows from (1.6), (1.7) and (2.9) (cf. [18, Sect. II.1.2]) that

$$\|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)} \le C \sup_{(\boldsymbol{0},0) \ne (\boldsymbol{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}\big((\boldsymbol{v}, q), (\boldsymbol{w}, r)\big)}{\|\boldsymbol{w}\|_{H^1(\Omega)} + \|r\|_{L_2(\Omega)}} \quad \forall\, (\boldsymbol{v}, q) \in V_h \times Q_h, \tag{2.10}$$

where the positive constant C depends only on the constants in (1.6), (1.7) and β_d. The estimate (2.10) implies that (2.8) is well-posed and, together with (2.1), also implies a quasi-optimal discretization error estimate [18, Sect. II.2.4]

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega)} + \|p - p_h\|_{L_2(\Omega)} \le C \left(\inf_{\boldsymbol{v} \in V_h} \|\boldsymbol{u} - \boldsymbol{v}\|_{H^1(\Omega)} + \inf_{q \in Q_h} \|p - q\|_{L_2(\Omega)} \right), \tag{2.11}$$

where the positive constant C depends only on the constants in (1.6), (1.7), (2.2) and (2.9).

We assume that the finite element spaces V_h and Q_h enjoy the following approximation properties:

$$\inf_{v \in V_h} \|\boldsymbol{\zeta} - \boldsymbol{v}\|_{H^1(\Omega)} \le Ch^\alpha \|\boldsymbol{\zeta}\|_{H^{1+\alpha}(\Omega)} \qquad \forall\, \boldsymbol{\zeta} \in [H_0^1(\Omega)]^d \cap [H^{1+\alpha}(\Omega)]^d, \tag{2.12}$$

$$\inf_{q \in Q_h} \|\mu - q\|_{L_2(\Omega)} \le Ch^\alpha \|\mu\|_{H^\alpha(\Omega)} \qquad \forall\, \mu \in L_2^0(\Omega) \cap H^\alpha(\Omega), \tag{2.13}$$

where the positive constant depends only on the shape regularity of $\mathcal{T}_h$. It then follows from (2.7), (2.11), (2.12) and (2.13) that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{H^1(\Omega)} + \|p - p_h\|_{L_2(\Omega)} \le Ch^\alpha \|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)}, \tag{2.14}$$

where the positive constant C depends on the shape regularity of $\mathcal{T}_h$ and the constants in (1.6), (1.7), (2.2) and (2.9).

The assumptions on V_h and Q_h are satisfied by many finite element pairs, such as the Taylor–Hood elements and the MINI element (cf. [5,18,24] and the references therein). For concreteness we will use the $P_2 - P_1$ Taylor–Hood element in the multigrid methods. But the results can also be extended to other finite element methods.

3 Multigrid methods

Let $\mathcal{T}_0$ be an initial triangulation of Ω and the triangulation $\mathcal{T}_k$ ($k \ge 1$) be generated from $\mathcal{T}_{k-1}$ through uniform refinement. The mesh parameter for $\mathcal{T}_k$ is denoted by h_k and we have $h_k = \frac{1}{2}h_{k-1}$. Let $V_k \subset [H_0^1(\Omega)]^d$ (resp. $Q_k \subset L_2^0(\Omega)$) be the continuous P_2 (resp. P_1) Lagrange finite element space associated with $\mathcal{T}_k$. Then $V_0 \times Q_0 \subset V_1 \times Q_1 \subset \cdots$ and the kth level finite element approximation $(\boldsymbol{u}_k, p_k) \in V_k \times Q_k$ for (1.3) is defined by

$$\mathcal{B}\big((\boldsymbol{u}_k, p_k), (\boldsymbol{v}, q)\big) = (\boldsymbol{f}, \boldsymbol{v}) \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \tag{3.1}$$

Let $\mathcal{N}_{k,1}$ (resp. $\mathcal{N}_{k,2}$) be the set of the nodes of the continuous P_1 (resp. P_2) finite element space associated with $\mathcal{T}_k$. The mesh-dependent inner products on V_k and Q_k are defined by

$$((\boldsymbol{v}, \boldsymbol{w}))_k = h_k^d \sum_{x \in \mathcal{N}_{k,2}} \boldsymbol{v}(x) \cdot \boldsymbol{w}(x) \qquad \forall\, \boldsymbol{v}, \boldsymbol{w} \in V_k, \tag{3.2a}$$

$$(q, r)_k = h_k^{d+2} \sum_{x \in \mathcal{N}_{k,1}} q(x) r(x) \qquad \forall\, q, r \in Q_k, \tag{3.2b}$$

and the mesh-dependent inner product $[\cdot, \cdot]_k$ on $V_k \times Q_k$ is given by

$$[(\boldsymbol{v}, q), (\boldsymbol{w}, r)]_k = ((\boldsymbol{v}, \boldsymbol{w}))_k + (q, r)_k. \tag{3.3}$$

Let the operator $\mathbb{B}_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ be defined by

$$[\mathbb{B}_k(\boldsymbol{w}, r), (\boldsymbol{v}, q)]_k = \mathcal{B}\big((\boldsymbol{w}, r), (\boldsymbol{v}, q)\big) \qquad \forall\, (\boldsymbol{v}, q), (\boldsymbol{w}, r) \in V_k \times Q_k. \tag{3.4}$$

Then $\mathbb{B}_k$ is nonsingular and also symmetric with respect to $[\cdot,\cdot]_k$. We can rewrite (3.1) as

$$\mathbb{B}_k(\boldsymbol{u}_k, p_k) = (\boldsymbol{f}_k, 0), \tag{3.5}$$

where $\boldsymbol{f}_k \in V_k$ is determined by $((\boldsymbol{f}_k, \boldsymbol{v}))_k = (\boldsymbol{f}, \boldsymbol{v})$ for all $\boldsymbol{v} \in V_k$.

Below we will construct multigrid methods for equations of the form

$$\mathbb{B}_k(\boldsymbol{v}, q) = (\boldsymbol{g}, z) \tag{3.6}$$

that includes (3.5) as a special case.

There are two main ingredients in multigrid methods. First we need intergrid transfer operators that move functions between grids. Since the finite element spaces are nested, we can take the coarse-to-fine operator $I_{k-1}^k : V_{k-1} \times Q_{k-1} \longrightarrow V_k \times Q_k$ to be the natural injection. Then we define the fine-to-coarse operator $I_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ to be the transpose of I_{k-1}^k with respect to the mesh-dependent inner products, i.e.,

$$[I_k^{k-1}(\boldsymbol{v}, q), (\boldsymbol{w}, r)]_{k-1} = [(\boldsymbol{v}, q), I_{k-1}^k(\boldsymbol{w}, r)]_k \tag{3.7}$$

for all $(\boldsymbol{v}, q) \in V_k \times Q_k$ and $(\boldsymbol{w}, r) \in V_{k-1} \times Q_{k-1}$.

Next we introduce two smoothers, which provide the second ingredient in multigrid methods.

3.1 Two smoothers

Let the discrete Laplace operator $\Delta_k : V_k \longrightarrow V_k$ be defined by

$$((-\Delta_k \boldsymbol{v}, \boldsymbol{w}))_k = \int_\Omega \nabla \boldsymbol{v} : \nabla \boldsymbol{w}\, dx \qquad \forall\, \boldsymbol{v}, \boldsymbol{w} \in V_k. \tag{3.8}$$

We take $L_k : V_k \longrightarrow V_k$ to be an optimal preconditioner for $-\Delta_k$ such that L_k is SPD with respect to $((\cdot,\cdot))_k$ and

$$\kappa_1 \le \lambda_{\min}\big(L_k(-\Delta_k)\big) \le \lambda_{\max}\big(L_k(-\Delta_k)\big) \le \kappa_2, \tag{3.9}$$

where κ_1 and κ_2 are positive constants independent of the mesh levels. Then we define the operator $\mathbb{S}_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ by

$$\mathbb{S}_k(\boldsymbol{v}, q) = (L_k \boldsymbol{v}, h_k^2 q). \tag{3.10}$$

It is clear that $\mathbb{S}_k$ is SPD with respect to $[\cdot,\cdot]_k$.

Remark 3.1 There are many choices for the optimal preconditioner L_k, including multigrid preconditioners [10] and domain decomposition preconditioners [38].

The first smoother for (3.6) is given by

$$(\boldsymbol{v}_{\text{new}}, q_{\text{new}}) = (\boldsymbol{v}_{\text{old}}, q_{\text{old}}) + \delta_k \mathbb{B}_k \mathbb{S}_k \big((\boldsymbol{g}, z) - \mathbb{B}_k(\boldsymbol{v}_{\text{old}}, q_{\text{old}})\big), \tag{3.11}$$

where $\delta_k > 0$ is a damping factor determined by the condition that the spectral radius $\rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k)$ satisfies

$$\rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k) \leq \frac{1}{\delta_k}. \tag{3.12}$$

This smoother will be used in the post-smoothing steps of the multigrid algorithms.

Remark 3.2 The smoother defined by (3.11) is Richardson relaxation for the equation

$$\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k(\boldsymbol{v}, q) = \mathbb{B}_k \mathbb{S}_k(\boldsymbol{g}, z)$$

that is equivalent to (3.6).

Remark 3.3 We can take the damping factor δ_k to be Ch_k^2, where the constant C is independent of the grid levels (cf. Corollary 4.3).

The second smoother for (3.6) is given by

$$(\boldsymbol{v}_{\text{new}}, q_{\text{new}}) = (\boldsymbol{v}_{\text{old}}, q_{\text{old}}) + \delta_k \mathbb{S}_k \mathbb{B}_k \big((\boldsymbol{g}, z) - \mathbb{B}_k(\boldsymbol{v}_{\text{old}}, q_{\text{old}})\big), \tag{3.13}$$

which will be used in the pre-smoothing steps of the multigrid algorithms.

Remark 3.4 The definition (3.13) of the pre-smoother is motivated by the fact that it is dual to the post-smoother with respect to the bilinear form $\mathcal{B}(\cdot, \cdot)$ (cf. (3.16)).

3.2 W-cycle and V-cycle multigrid algorithms

Let the output of the W-cycle algorithm for (3.6) with initial guess $(\boldsymbol{v}_0, q_0) \in V_k \times Q_k$ and m_1 (resp. m_2) pre-smoothing (resp. post-smoothing) steps be denoted by $MG_W(k, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2)$. For $k = 0$, we take $MG_W(0, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2)$ to be $\mathbb{B}_0^{-1}(\boldsymbol{g}, z)$. For $k \geq 1$, we compute $MG_W(k, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2)$ recursively in three steps.

- *Pre-smoothing* Apply the iteration defined by (3.13) m_1 times with initial guess $(\boldsymbol{v}_0, q_0)$ to obtain $(\boldsymbol{v}_{m_1}, q_{m_1})$.
- *Coarse grid correction* Let $(\boldsymbol{g}', z') = I_k^{k-1}\big((\boldsymbol{g}, z) - \mathbb{B}_k(\boldsymbol{v}_{m_1}, q_{m_1})\big)$ be the transferred residual of $(\boldsymbol{v}_{m_1}, q_{m_1})$. We compute $(\boldsymbol{v}_1', q_1'), (\boldsymbol{v}_2', q_2') \in V_{k-1} \times Q_{k-1}$ by

$$(\boldsymbol{v}_1', q_1') = MG_W(k-1, (\boldsymbol{g}', z'), (\boldsymbol{0}, 0), m_1, m_2),$$
$$(\boldsymbol{v}_2', q_2') = MG_W(k-1, (\boldsymbol{g}', z'), (\boldsymbol{v}_1', q_1'), m_1, m_2),$$

  and take $(\boldsymbol{v}_{m_1+1}, q_{m_1+1})$ to be $(\boldsymbol{v}_{m_1}, q_{m_1}) + I_{k-1}^k(\boldsymbol{v}_2', q_2')$.

- *Post-smoothing* Apply the iteration defined by (3.11) m_2 times with initial guess $(\boldsymbol{v}_{m_1+1}, q_{m_1+1})$ to obtain $(\boldsymbol{v}_{m_1+m_2+1}, q_{m_1+m_2+1})$.

The final output is $MG_W(k, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2) = (\boldsymbol{v}_{m_1+m_2+1}, q_{m_1+m_2+1})$.

We denote by $MG_V(k, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2)$ the output of the V-cycle algorithm for (3.6) with initial guess $(\boldsymbol{v}_0, q_0) \in V_k \times Q_k$ and m_1 (resp. m_2) pre-smoothing (resp. post-smoothing) steps. The computation of $MG_V(k, (\boldsymbol{g}, z), (\boldsymbol{v}_0, q_0), m_1, m_2)$ is identical with the computation for the W-cycle algorithm except in the coarse grid correction step, where we compute

$$(\boldsymbol{v}_1', q_1') = MG_V(k-1, (\boldsymbol{g}', z'), (\boldsymbol{0}, 0), m_1, m_2)$$

and take$(\boldsymbol{v}_{m_1+1}, q_{m_1+1})$ to be $(\boldsymbol{v}_{m_1}, q_{m_1}) + I_{k-1}^k(\boldsymbol{v}_1', q_1')$.

Remark 3.5 In this paper we will only establish the uniform convergence of the W-cycle algorithm (cf. Sect. 5). But numerical results (cf. Sect. 6) indicate that the V-cycle algorithm is also uniformly convergent.

3.3 Error propagation operators

The effect of one post-smoothing step defined by (3.11) is measured by the operator

$$R_k = Id_k - \delta_k \mathbb{B}_k \mathbb{S}_k \mathbb{B}_k, \tag{3.14}$$

where Id_k is the identity operator on $V_k \times Q_k$. On the other hand the effect of one pre-smoothing step defined by (3.13) is measured by the operator

$$S_k = Id_k - \delta_k \mathbb{S}_k \mathbb{B}_k^2. \tag{3.15}$$

The following relation between the two operators R_k and S_k is a simple consequence of (3.4), (3.14) and (3.15):

$$\mathcal{B}\big(R_k(\boldsymbol{v}, q), (\boldsymbol{w}, r)\big) = \mathcal{B}\big((\boldsymbol{v}, q), S_k(\boldsymbol{w}, r)\big) \qquad \forall (\boldsymbol{v}, q), (\boldsymbol{w}, r) \in V_k \times Q_k. \tag{3.16}$$

The error propagation operator $E_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ for the multigrid algorithms satisfies a well-known recursive relation:

$$E_k = R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1} + I_{k-1}^k E_{k-1}^p P_k^{k-1}) S_k^{m_1}, \tag{3.17}$$

where $p = 2$ (resp. $p = 1$) for the W-cycle (resp. V-cycle) algorithm, and $P_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ is the transpose of I_{k-1}^k with respect to the bilinear form $\mathcal{B}(\cdot, \cdot)$, i.e.,

$$\mathcal{B}\big(P_k^{k-1}(\boldsymbol{v}, q), (\boldsymbol{w}, r)\big) = \mathcal{B}\big((\boldsymbol{v}, q), I_{k-1}^k(\boldsymbol{w}, r)\big) \tag{3.18}$$

for all $(\boldsymbol{v}, q) \in V_k \times Q_k$ and $(\boldsymbol{w}, r) \in V_{k-1} \times Q_{k-1}$. By construction we also have $E_0 = 0$.

Since I_{k-1}^k is the natural injection, the operator P_k^{k-1} is just the restriction to $V_k \times Q_k$ of the Ritz projection from $[H_0^1(\Omega)]^d \times L_2^0(\Omega)$ to $V_{k-1} \times Q_{k-1}$. In particular, we have

$$P_k^{k-1} I_{k-1}^k = Id_{k-1}, \quad (Id_k - I_{k-1}^k P_k^{k-1})^2 = Id_k - I_{k-1}^k P_k^{k-1}, \tag{3.19}$$

and the Galerkin orthogonality

$$0 = \mathcal{B}\big((Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q), I_{k-1}^k(\boldsymbol{w}, r)\big) \tag{3.20}$$

that holds for all $(\boldsymbol{v}, q) \in V_k \times Q_k$ and $(\boldsymbol{w}, r) \in V_{k-1} \times Q_{k-1}$.

4 Mesh-dependent norms

In this section we introduce mesh-dependent norms which are useful tools for the convergence analysis of the W-cycle algorithm in Sect. 5. In order to avoid the proliferation of constants, from now on we will use $A \lesssim B$ (or $B \gtrsim A$) to represent the inequality $A \leq (\text{constant})B$, where the positive constant is independent of the grid size and the grid level. The statement $A \approx B$ is equivalent to $A \lesssim B$ and $B \lesssim A$.

For $0 \leq s \leq 1$, we use the SPD operator $\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k$ to define the first family of mesh-dependent norms $\|\cdot\|_{s,k}$ by

$$\|(\boldsymbol{v}, q)\|_{s,k}^2 = [(\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^s(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.1}$$

It follows immediately from (3.2) and standard discrete estimates [16,19] that

$$((\boldsymbol{v}, \boldsymbol{v}))_k \approx \|\boldsymbol{v}\|_{L_2(\Omega)}^2 \qquad \forall\, \boldsymbol{v} \in V_k, \tag{4.2a}$$

$$(q, q)_k \approx h_k^2 \|q\|_{L_2(\Omega)}^2 \qquad \forall\, q \in Q_k. \tag{4.2b}$$

The estimates in (4.2) and the definitions (3.3) and (4.1) imply

$$\|(\boldsymbol{v}, q)\|_{0,k}^2 = [(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k \approx \|\boldsymbol{v}\|_{L_2(\Omega)}^2 + h_k^2 \|q\|_{L_2(\Omega)}^2 \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.3}$$

Lemma 4.1 *We have*

$$\|(\boldsymbol{v}, q)\|_{1,k} \approx \|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.4}$$

Proof From (2.1) and (2.10) we have

$$\|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)} \approx \sup_{(\boldsymbol{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{\mathcal{B}\big((\boldsymbol{v}, q), (\boldsymbol{w}, r)\big)}{\|\boldsymbol{w}\|_{H^1(\Omega)} + \|r\|_{L_2(\Omega)}} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.5}$$

By construction (cf. (3.8) and (3.9)) we also have

$$((L_k^{-1}\boldsymbol{w}, \boldsymbol{w}))_k \approx (((-\Delta_k)\boldsymbol{w}, \boldsymbol{w}))_k = |\boldsymbol{w}|_{H^1(\Omega)}^2 \approx \|\boldsymbol{w}\|_{H^1(\Omega)}^2 \qquad \forall\, \boldsymbol{w} \in V_k,$$

which together with (3.3), (3.10) and (4.2b) implies

$$\begin{aligned}[\mathbb{S}_k^{-1}(\boldsymbol{w}, r), (\boldsymbol{w}, r)]_k &= ((L_k^{-1}\boldsymbol{w}, \boldsymbol{w}))_k + h_k^{-2}(r, r)_k \\ &\approx \|\boldsymbol{w}\|^2_{H^1(\Omega)} + \|r\|^2_{L_2(\Omega)} \qquad \forall\,(\boldsymbol{w}, r) \in V_k \times Q_k. \end{aligned} \tag{4.6}$$

Let $(\boldsymbol{v}, q) \in V_k \times Q_k$ be arbitrary and $(\boldsymbol{z}, g) = \mathbb{S}_k\mathbb{B}_k(\boldsymbol{v}, q)$. It follows from (3.4), (4.1), (4.5), (4.6) and duality that

$$\begin{aligned}\|(\boldsymbol{v}, q)\|_{1,k} &= [\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k^{\frac{1}{2}} \\ &= [\mathbb{S}_k^{-1}(\mathbb{S}_k\mathbb{B}_k)(\boldsymbol{v}, q), (\mathbb{S}_k\mathbb{B}_k)(\boldsymbol{v}, q)]_k^{\frac{1}{2}} \\ &= [\mathbb{S}_k^{-1}(\boldsymbol{z}, g), (\boldsymbol{z}, g)]_k^{\frac{1}{2}} \\ &= \sup_{(\boldsymbol{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{[\mathbb{S}_k^{-1}(\boldsymbol{z}, g), (\boldsymbol{w}, r)]_k}{[\mathbb{S}_k^{-1}(\boldsymbol{w}, r), (\boldsymbol{w}, r)]_k^{\frac{1}{2}}} \\ &\approx \sup_{(\boldsymbol{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{[\mathbb{B}_k(\boldsymbol{v}, q), (\boldsymbol{w}, r)]_k}{\|\boldsymbol{w}\|_{H^1(\Omega)} + \|r\|_{L_2(\Omega)}} \\ &= \sup_{(\boldsymbol{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{\mathcal{B}\big((\boldsymbol{v}, q), (\boldsymbol{w}, r)\big)}{\|\boldsymbol{w}\|_{H^1(\Omega)} + \|r\|_{L_2(\Omega)}} \approx \|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)}.\end{aligned}$$

□

Corollary 4.2 *We have*

$$\|P_k^{k-1}(\boldsymbol{v}, q)\|_{1,k-1} \lesssim \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.7}$$

Proof Since I_{k-1}^k is the natural injection, we have an obvious estimate

$$\|I_{k-1}^k(\boldsymbol{w}, r)\|_{1,k} \approx \|(\boldsymbol{w}, r)\|_{1,k-1} \qquad \forall\,(\boldsymbol{w}, r) \in V_{k-1} \times Q_{k-1} \tag{4.8}$$

because of (4.4).

The estimate (4.7) follows from (2.1), (3.18), (4.4), (4.5) and (4.8). □

Corollary 4.3 *The spectral radius of the operator $\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k$ is bounded by Ch_k^{-2}, where the positive constant C is independent of grid sizes and grid levels.*

Proof It follows from (4.1), (4.3), (4.4) and the Rayleigh quotient formula that

$$\begin{aligned}\lambda_{\max}(\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k) &= \max_{(\boldsymbol{0},0)\notin(\boldsymbol{v},q)} \frac{[\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k}{[(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k} \\ &= \max_{(\boldsymbol{0},0)\notin(\boldsymbol{v},q)} \frac{\|(\boldsymbol{v}, q)\|^2_{1,k}}{\|(\boldsymbol{v}, q)\|^2_{0,k}} \\ &\leq C \max_{(\boldsymbol{0},0)\notin(\boldsymbol{v},q)} \frac{\|\boldsymbol{v}\|^2_{H^1(\Omega)} + \|q\|^2_{L_2(\Omega)}}{\|\boldsymbol{v}\|^2_{L_2(\Omega)} + h_k^2\|q\|^2_{L_2(\Omega)}} \leq Ch_k^{-2},\end{aligned}$$

where the last inequality comes from a standard inverse estimate [16,19]. □

The next result provides a link between a mesh-dependent norm and certain Sobolev norms.

Lemma 4.4 *We have*

$$\|(\boldsymbol{v}, q)\|_{1-\alpha,k} \approx \|\boldsymbol{v}\|_{H^{1-\alpha}(\Omega)} + h_k^\alpha \|q\|_{L_2(\Omega)} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k, \tag{4.9}$$

where $\alpha \in \left(\frac{1}{2}, 1\right]$ *is the index of elliptic regularity in* (2.7).

Proof Since $V_k \times Q_k$ is a subspace of $[H_0^1(\Omega)]^d \times L_2^0(\Omega)$, we have

$$\|\boldsymbol{v}\|_{H^{1-\alpha}(\Omega)} + h_k^\alpha \|q\|_{L_2(\Omega)} \lesssim \|(\boldsymbol{v}, q)\|_{1-\alpha,k} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k$$

by (4.1), (4.3), (4.4) and interpolation [37,39].

Let $\Pi_k : [L_2(\Omega)]^d \longrightarrow V_k$ be the L_2 orthogonal projection. It is well-known (cf. [9]) that

$$\|\Pi_k \boldsymbol{\zeta}\|_{H^1(\Omega)} \lesssim \|\boldsymbol{\zeta}\|_{H^1(\Omega)} \qquad \forall\,\boldsymbol{\zeta} \in [H_0^1(\Omega)]^d,$$

which implies by (4.1), (4.3), (4.4) and interpolation

$$\|(\Pi_k \boldsymbol{\zeta}, 0)\|_{1-\alpha,k} \lesssim \|\boldsymbol{\zeta}\|_{H^{1-\alpha}(\Omega)} \qquad \forall\,\boldsymbol{\zeta} \in [H^{1-\alpha}(\Omega)]^d. \tag{4.10}$$

Similarly we have

$$\|(\mathbf{0}, \pi_k \xi)\|_{1-\alpha,k} \lesssim h_k^\alpha \|\xi\|_{L_2(\Omega)} \qquad \forall\,\xi \in L_2^0(\Omega), \tag{4.11}$$

where $\pi_k : L_2^0(\Omega) \longrightarrow Q_k$ is the L_2 orthogonal projection.

Applying (4.10) to $\boldsymbol{\zeta} = \boldsymbol{v}$ and (4.11) to $\xi = q$ we arrive at the estimate in the other direction:

$$\|(\boldsymbol{v}, q)\|_{1-\alpha,k} \lesssim \|\boldsymbol{v}\|_{H^{1-\alpha}(\Omega)} + h_k^\alpha \|q\|_{L_2(\Omega)} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k.$$

□

For $1 \le s \le 2$, we define a second family of mesh-dependent norms $|||\cdot|||_{s,k}$ by duality:

$$|||(\boldsymbol{v}, q)|||_{s,k} = \sup_{(\mathbf{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{\mathcal{B}\big((\boldsymbol{v}, q), (\boldsymbol{w}, r)\big)}{\|(\boldsymbol{w}, r)\|_{2-s,k}} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.12}$$

It follows immediately from (4.4), (4.5) and (4.12) that

$$|||(\boldsymbol{v}, q)|||_{1,k} \approx \|\boldsymbol{v}\|_{H^1(\Omega)} + \|q\|_{L_2(\Omega)} \approx \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\,(\boldsymbol{v}, q) \in V_k \times Q_k. \tag{4.13}$$

Note also that (4.12) implies

$$\mathcal{B}\big((\boldsymbol{v}, q), (\boldsymbol{w}, r)\big) \le |||(\boldsymbol{v}, q)|||_{s,k} \|(\boldsymbol{w}, r)\|_{2-s,k} \quad \forall\,(\boldsymbol{v}, q), (\boldsymbol{w}, r) \in V_k \times Q_k,\ 1 \le s \le 2. \tag{4.14}$$

5 Convergence analysis for the W-cycle algorithm

We follow the classical approach in [3] for the analysis of the W-cycle algorithm, which is based on the smoothing and approximation properties.

5.1 Smoothing property

We have a standard smoothing property for R_k since the damping factor $\delta_k = Ch_k^2$ (cf. Remark 3.3) satisfies condition (3.12).

Lemma 5.1 *The following estimate holds for the operator* R_k :

$$\|R_k^m(\boldsymbol{v}, q)\|_{1,k} \lesssim h_k^{-\tau} m^{-\tau/2} \|(\boldsymbol{v}, q)\|_{1-\tau,k} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k,\ 0 \le \tau \le 1. \quad (5.1)$$

Proof Let $(\boldsymbol{v}, q) \in V_k \times Q_k$ be arbitrary. It follows from (3.12), (3.14), (4.1) and the spectral theorem that

$$\begin{aligned}
\|R_k^m(\boldsymbol{v}, q)\|_{1,k}^2 &= [\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k(Id_k - \delta_k\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^m(\boldsymbol{v}, q), (Id_k - \delta_k\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^m(\boldsymbol{v}, q)]_k \\
&= \delta_k^{-\tau}[(\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^{1-\tau}(\delta_k\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^{\tau}(Id_k - \delta_k\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^m(\boldsymbol{v}, q), (Id_k - \delta_k\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^m(\boldsymbol{v}, q)]_k \\
&\lesssim h_k^{-2\tau} \max_{0\le x\le 1}[(1-x)^{2m}x^{\tau}][(\mathbb{B}_k\mathbb{S}_k\mathbb{B}_k)^{1-\tau}(\boldsymbol{v}, q), (\boldsymbol{v}, q)]_k \\
&\lesssim h_k^{-2\tau} m^{-\tau} \|(\boldsymbol{v}, q)\|_{1-\tau,k}^2.
\end{aligned}$$

□

Remark 5.2 In the special case where $\tau = 0$, the arguments in the proof of Lemma 5.1 lead to the following estimate:

$$\|R_k^m(\boldsymbol{v}, q)\|_{1,k} \le \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k. \quad (5.2)$$

Remark 5.3 The smoothing property (5.1) for R_k implies a corresponding smoothing property for the operator S_k through duality. But we will only need the estimate

$$\|S_k(\boldsymbol{v}, q)\|_{1,k} \lesssim \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k \quad (5.3)$$

that follows immediately from (3.16), (4.12), (4.13) and (5.2).

5.2 Approximation property

Recall $\alpha \in \left(\frac{1}{2}, 1\right]$ is the index of elliptic regularity in (2.7).

Lemma 5.4 *We have, for* $k \ge 1$,

$$\|(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)\|_{1-\alpha,k} \lesssim h_k^{\alpha} \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\, (\boldsymbol{v}, q) \in V_k \times Q_k.$$

Proof Let $(\boldsymbol{v}, q) \in V_k \times Q_k$ be arbitrary and $(\boldsymbol{\zeta}, \mu) = (Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)$. In view of (4.9), it suffices to show that

$$\|\boldsymbol{\zeta}\|_{H^{1-\alpha}(\Omega)} + h_k^\alpha \|\mu\|_{L_2(\Omega)} \lesssim h_k^\alpha \|(\boldsymbol{v}, q)\|_{1,k}.$$

The estimate for μ follows immediately from (4.4), (4.7) and (4.8):

$$\|\mu\|_{L_2(\Omega)} \lesssim \|(\boldsymbol{\zeta}, \mu)\|_{1,k} = \|(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)\|_{1,k} \lesssim \|(\boldsymbol{v}, q)\|_{1,k}.$$

We will prove the estimate for $\boldsymbol{\zeta}$ by a duality argument.

Let $\boldsymbol{\chi} \in [H^{-1+\alpha}(\Omega)]^d$ be arbitrary, $(\boldsymbol{\xi}, \theta) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$ satisfy

$$\mathcal{B}\big((\boldsymbol{\xi}, \theta), (\boldsymbol{w}, r)\big) = \langle \boldsymbol{\chi}, \boldsymbol{w} \rangle \qquad \forall\, (\boldsymbol{w}, r) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega), \tag{5.4}$$

and $(\boldsymbol{\xi}_{k-1}, \theta_{k-1}) \in V_{k-1} \times Q_{k-1}$ satisfy

$$\mathcal{B}\big((\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (\boldsymbol{w}, r)\big) = \langle \boldsymbol{\chi}, \boldsymbol{w} \rangle \qquad \forall\, (\boldsymbol{w}, r) \in V_{k-1} \times Q_{k-1}. \tag{5.5}$$

It follows from (2.1), (3.20), (4.4), (5.4) and (5.5) that

$$\begin{aligned}
\langle \boldsymbol{\chi}, \boldsymbol{\zeta} \rangle &= \mathcal{B}\big((\boldsymbol{\xi}, \theta), (\boldsymbol{\zeta}, \mu)\big) \\
&= \mathcal{B}\big((\boldsymbol{\xi}, \theta), (Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)\big) \\
&= \mathcal{B}\big((\boldsymbol{\xi}, \theta) - (\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)\big) \\
&= \mathcal{B}\big((\boldsymbol{\xi}, \theta) - (\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (\boldsymbol{v}, q)\big) \\
&\lesssim (\|\boldsymbol{\xi} - \boldsymbol{\xi}_{k-1}\|_{H^1(\Omega)} + \|\theta - \theta_{k-1}\|_{L_2(\Omega)}) \|(\boldsymbol{v}, q)\|_{1,k},
\end{aligned}$$

which together with the discretization error estimate (2.14) yields

$$\langle \boldsymbol{\chi}, \boldsymbol{\zeta} \rangle \lesssim h_k^\alpha \|\boldsymbol{\chi}\|_{H^{-1+\alpha}(\Omega)} \|(\boldsymbol{v}, q)\|_{1,k} \qquad \forall\, \boldsymbol{\chi} \in [H^{-1+\alpha}(\Omega)]^d. \tag{5.6}$$

The estimate for $\boldsymbol{\zeta}$ follows from (5.6) and duality. □

5.3 Convergence of the W-cycle algorithm

We begin with the two-grid algorithm where the coarse grid residual equation is solved exactly. In view of (3.17), the error propagation operator for the two-grid algorithm is given by $R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1}) S_k^{m_1}$. We will divide the analysis of the two-grid algorithm into three cases.

In the first case we assume $m_1 = 0$ and $m_2 = m$, i.e., we have a one-sided algorithm with only post-smoothing. It follows from Lemmas 5.1 and 5.4 that

$$\begin{aligned}\|R_k^m(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v},q)\|_{1,k} &\lesssim (h_k^{-\alpha} m^{-\alpha/2})\|(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v},q)\|_{1-\alpha,k}\\ &\lesssim (h_k^{-\alpha} m^{-\alpha/2}) h_k^{\alpha}\|(\boldsymbol{v},q)\|_{1,k}\\ &= m^{-\alpha/2}\|(\boldsymbol{v},q)\|_{1,k} \qquad \forall (\boldsymbol{v},q)\in V_k\times Q_k.\end{aligned} \tag{5.7}$$

In the second case we assume $m_2 = 0$ and $m_1 = m$, i.e., we have a one-sided algorithm with only pre-smoothing. From (3.16), (3.18), (4.12), (4.14) and (5.7) we find

$$\begin{aligned}&|||(Id_k - I_{k-1}^k P_k^{k-1})S_k^m(\boldsymbol{v},q)|||_{1,k}\\ &= \sup_{(\mathbf{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{\mathcal{B}\big((Id_k - I_{k-1}^k P_k^{k-1})S_k^m(\boldsymbol{v},q),(\boldsymbol{w},r)\big)}{\|(\boldsymbol{w},r)\|_{1,k}}\\ &= \sup_{(\mathbf{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{\mathcal{B}\big((\boldsymbol{v},q),R_k^m(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{w},r)\big)}{\|(\boldsymbol{w},r)\|_{1,k}}\\ &\le \sup_{(\mathbf{0},0)\neq(\boldsymbol{w},r)\in V_k\times Q_k} \frac{|||(\boldsymbol{v},q)|||_{1,k}\|R_k^m(Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{w},r)\|_{1,k}}{\|(\boldsymbol{w},r)\|_{1,k}}\\ &\lesssim m^{-\alpha/2}|||(\boldsymbol{v},q)|||_{1,k} \qquad \forall(\boldsymbol{v},q)\in V_k\times Q_k.\end{aligned} \tag{5.8}$$

In the last case we assume $m_1, m_2 \ge 1$. From (3.18), (3.19), (4.13), (5.7) and (5.8) we have

$$\begin{aligned}&\|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\boldsymbol{v},q)\|_{1,k}\\ &= \|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\boldsymbol{v},q)\|_{1,k}\\ &\lesssim m_2^{-\alpha/2}\|(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\boldsymbol{v},q)\|_{1,k}\\ &\approx m_2^{-\alpha/2}|||(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\boldsymbol{v},q)|||_{1,k}\\ &\lesssim m_2^{-\alpha/2}m_1^{-\alpha/2}|||(\boldsymbol{v},q)|||_{1,k} \approx (m_1m_2)^{-\alpha/2}\|(\boldsymbol{v},q)\|_{1,k} \qquad \forall(\boldsymbol{v},q)\in V_k\times Q_k.\end{aligned} \tag{5.9}$$

Combining (4.13) and (5.7), (5.8), (5.9), we have the following estimate for the two-grid algorithm:

$$\begin{aligned}&\|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\boldsymbol{v},q)\|_{1,k}\\ &\le C_*\big(\max(1,m_1)\max(1,m_2)\big)^{-\alpha/2}\|(\boldsymbol{v},q)\|_{1,k}\end{aligned} \tag{5.10}$$

for all $(\boldsymbol{v},q)\in V_k\times Q_k$, where the positive constant C_* is independent of k.

We can now use the two-grid estimate (5.10) and a perturbation argument to analyze the W-cycle algorithm.

Theorem 5.5 *Let E_k be the error propagation operator for the kth level W-cycle algorithm. For any $C_\dagger > C_*$ (the constant in (5.10)), there exists a positive number m_* (independent of k) such that*

$$\|E_k(\boldsymbol{v}, q)\|_{1,k} \le C_\dagger \big(\max(1, m_1)\max(1, m_2)\big)^{-\alpha/2} \|(\boldsymbol{v}, q)\|_{1,k} \tag{5.11}$$

for all $(\boldsymbol{v}, q) \in V_k \times Q_k$ and $k \ge 0$, provided $\max(1, m_1)\max(1, m_2) \ge m_*$.

Proof We will derive (5.11) by mathematical induction. The estimate (5.11) is trivially true for $k = 0$ since $E_0 = 0$.

Let $m = \max(1, m_1)\max(1, m_2)$. Suppose (5.11) is valid for $k = j - 1$. From (4.7), (4.8), (5.2), (5.3) and the induction hypothesis, we have

$$\|R_j^{m_2} I_{j-1}^j E_{j-1}^2 P_j^{j-1} S_j^{m_1}(\boldsymbol{v}, q)\|_{1,j} \le C_\sharp C_\dagger^2 m^{-\alpha} \|(\boldsymbol{v}, q)\|_{1,j}$$

for all $(\boldsymbol{v}, q) \in V_j \times Q_j$, where the positive constant $C_\sharp$ is independent of j. The relation (3.17) for $p = 2$ and the estimate (5.10) imply

$$\begin{aligned}\|E_j(\boldsymbol{v}, q)\|_{1,j} &= \|R_j^{m_2}(Id_j - I_{j-1}^j P_j^{j-1}) S_j^{m_1}(\boldsymbol{v}, q) + R_j^{m_2} I_{j-1}^j E_{j-1}^2 P_j^{j-1} S_j^{m_1}(\boldsymbol{v}, q)\|_{1,j} \\ &\le \big[C_* m^{-\alpha/2} + C_\sharp C_\dagger^2 m^{-\alpha}\big] \|(\boldsymbol{v}, q)\|_{1,j}.\end{aligned}$$

Let $m_* > 0$ satisfy

$$C_\sharp C_\dagger^2 m_*^{-\alpha/2} \le C_\dagger - C_*.$$

Then, for $m \ge m_*$, we have

$$C_* m^{-\alpha/2} + C_\sharp C_\dagger^2 m^{-\alpha} \le m^{-\alpha/2}(C_* + C_\sharp C_\dagger^2 m_*^{-\alpha/2}) \le C_\dagger m^{-\alpha/2},$$

which completes the proof. □

Remark 5.6 The constant C_* in (5.10) depends only on the shape regularity of the triangulation $\mathcal{T}_0$ and the constants in (1.6), (1.7), (2.2), (2.7), (2.9) and (3.9). This also holds for the constants $C_\dagger$ and m_* in Theorem 5.5 if for example we take $C_\dagger$ to be $2C_*$. For the two dimensional Lamé problem, Proposition 2.1 implies that the estimates (5.10) and (5.11) are robust with respect to the Lamé constant λ.

We can draw the following conclusion from (4.4) and (5.11). If $\max(1, m_1)\max(1, m_2)$ (independent of k) is sufficiently large, then the W-cycle algorithm for the saddle point problem (1.3) is uniformly convergent with respect to the energy norm, i.e., the W-cycle algorithm is a contraction with respect to the energy norm and the contraction number is bounded away from 1 for all k. In particular this result holds for the Stokes problem (1.1) and the Lamé problem (1.2) discretized by the P_2-P_1 Taylor–Hood element.

Remark 5.7 With a slight modification of the smoothing precess we can also establish $O(1/m^{\alpha})$ contraction number estimates for one-sided algorithms with respect to appropriate mesh-dependent norms.

If in the one-sided algorithm with only pre-smoothing we apply m smoothing steps defined by (3.11) followed by m smoothing steps defined by (3.13), then for the corresponding two-grid algorithm we have the estimate

$$\begin{aligned}
&\|(Id_k - I_{k-1}^k P_k^{k-1}) S_k^m R_k^m (\boldsymbol{v}, q)\|_{1-\alpha,k} \\
&\quad = \|(Id_k - I_{k-1}^k P_k^{k-1})(Id_k - I_{k-1}^k P_k^{k-1}) S_k^m R_k^m (\boldsymbol{v}, q)\|_{1-\alpha,k} \\
&\quad \lesssim h_k^{\alpha} \|(Id_k - I_{k-1}^k P_k^{k-1}) S_k^m R_k^m (\boldsymbol{v}, q)\|_{1,k} \\
&\quad \approx h_k^{\alpha} |\!|\!|(Id_k - I_{k-1}^k P_k^{k-1}) S_k^m R_k^m (\boldsymbol{v}, q)|\!|\!|_{1,k} \\
&\quad \lesssim h_k^{\alpha} m^{-\alpha/2} |\!|\!| R_k^m (\boldsymbol{v}, q)|\!|\!|_{1,k} \\
&\quad \lesssim h_k^{\alpha} m^{-\alpha/2} \| R_k^m (\boldsymbol{v}, q)\|_{1,k} \\
&\quad \lesssim m^{-\alpha} \|(\boldsymbol{v}, q)\|_{1-\alpha,k} \qquad \forall (\boldsymbol{v}, q) \in V_k \times Q_k
\end{aligned} \tag{5.12}$$

that follows from (3.19), (4.13), Lemmas 5.1, 5.4 and (5.8).

If in the one-sided algorithm with only post-smoothing we apply m smoothing steps defined by (3.13) followed by m smoothing steps defined by (3.11), then the estimate for the corresponding two-grid algorithm is given by

$$|\!|\!| R_k^m S_k^m (Id_k - I_{k-1}^k P_k^{k-1})(\boldsymbol{v}, q)|\!|\!|_{1+\alpha,k} \lesssim m^{-\alpha} |\!|\!|(\boldsymbol{v}, q)|\!|\!|_{1+\alpha,k} \qquad \forall (\boldsymbol{v}, q) \in V_k \times Q_k, \tag{5.13}$$

which follows from (3.16), (3.18), (4.12), (4.14) and (5.12).

Uniform convergence of the W-cycle in the norm $\|\cdot\|_{1-\alpha,k}$ (resp. $|\!|\!|\cdot|\!|\!|_{1+\alpha,k}$) for the one-sided algorithm with only pre-smoothing (resp. post-smoothing) follows from (5.12) (resp. (5.13)) and arguments similar to the ones in the proof of Theorem 5.5.

6 Numerical results

In this section we report numerical results for the proposed W-cycle and V-cycle multigrid methods for the Stokes problem and the Lamé problem on two dimensional domains. In these tests we use the P_2-P_1 Taylor–Hood finite element in the discretization. The operator L_k is generated by the multigrid V(2,2) algorithm for the Laplace operator.

We first test the W-cycle algorithm for the Stokes problem on the unit square. Note that in this case we can take the index of elliptic regularity α to be 1. The initial triangulation $\mathcal{T}_0$ consists of eight triangles of the same size (Fig. 1). For $k \geq 1$, the kth level triangulation $\mathcal{T}_k$ is obtained from $\mathcal{T}_{k-1}$ by a uniform refinement.

The contraction numbers of the W-cycle algorithm are listed in Table 1. The contraction numbers are computed as the largest eigenvalue of the error propagation operator E_k. The leftmost column displays the numbers (m_1, m_2) of smoothing steps in the algorithm, whereas contraction numbers associated with the kth level triangulation are displayed in the other columns for various k. This table clearly

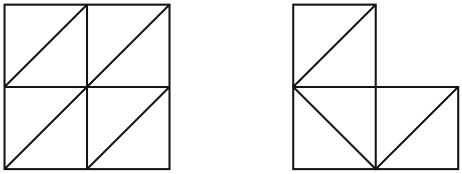

Fig. 1 The initial triangulation of the unit square (*left*); the initial triangulation of the L-shaped domain (*right*)

Table 1 Contraction numbers of the W-cycle algorithm for the Stokes problem on the unit square

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.82	0.86	0.86	0.86	0.86	0.86
(2, 2)	0.76	0.78	0.78	0.78	0.78	0.78
(4, 4)	0.66	0.68	0.69	0.69	0.69	0.69
(8, 8)	0.55	0.56	0.56	0.56	0.56	0.56
(16, 16)	0.38	0.39	0.39	0.39	0.39	0.39
(32, 32)	0.19	0.19	0.19	0.19	0.19	0.19

Table 2 Contraction numbers of the V-cycle algorithm for the Stokes problem on the unit square

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.87	0.90	0.90	0.90	0.90	0.90
(2, 2)	0.76	0.76	0.79	0.80	0.80	0.80
(4, 4)	0.66	0.70	0.70	0.70	0.70	0.70
(8, 8)	0.55	0.57	0.57	0.58	0.58	0.58
(16, 16)	0.38	0.40	0.40	0.40	0.40	0.40
(32, 32)	0.19	0.19	0.20	0.20	0.20	0.20

shows that with our new smoothers, the proposed W-cycle iterations are contractions, with contraction numbers independent of the grid level k. Furthermore, the asymptotic rate of decrease $(m_1 m_2)^{-1/2}$ for the contraction numbers is observed in these tests when the number of smoothing steps increases, which agrees with Theorem 5.5.

The contraction numbers of the V-cycle algorithm for the same problem are presented in Table 2. It can be observed that, similar to the W-cycle iterations, the multigrid V-cycle iterations converge uniformly on all grid levels for the Stokes problem. The contraction numbers in Table 2 exhibit a similar asymptotic rate of decrease as $m_1 m_2$ increases, but they are in general larger than the corresponding numbers in Table 1.

We also test the W-cycle algorithm on an L-shaped domain, where the kth level triangulation $\mathcal{T}_k$ $(k \geq 1)$ is obtained by k successive uniform refinements of the initial triangulation $\mathcal{T}_0$ with six triangles (Fig. 1). From the contraction numbers in Table 3 it is clear that, as predicted by Theorem 5.5, the W-cycle iterations converge uniformly, independent of the grid level k. It is also observed that the convergence rate is worse than the rate for the unit square, reflecting the fact that the index of elliptic regularity α is strictly less than 1 for the L-shaped domain.

Table 3 Contraction numbers of the W-cycle algorithm for the Stokes problem on the L-shaped domain

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.88	0.90	0.90	0.90	0.90	0.90
(2, 2)	0.81	0.83	0.84	0.84	0.84	0.84
(4, 4)	0.73	0.75	0.75	0.75	0.75	0.75
(8, 8)	0.63	0.65	0.65	0.65	0.65	0.65
(16, 16)	0.48	0.49	0.49	0.49	0.49	0.49
(32, 32)	0.28	0.29	0.29	0.29	0.29	0.29

Table 4 Contraction numbers of the W-cycle algorithm for the Lamé problem on the unit square with $\mu = 1$ and $\lambda = 500$

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.92	0.93	0.93	0.93	0.93	0.93
(2, 2)	0.93	0.93	0.93	0.93	0.93	0.93
(4, 4)	0.88	0.89	0.90	0.90	0.90	0.90
(8, 8)	0.81	0.83	0.83	0.83	0.83	0.83
(16, 16)	0.70	0.72	0.72	0.72	0.72	0.72
(32, 32)	0.53	0.54	0.54	0.54	0.54	0.54
(64, 64)	0.31	0.32	0.32	0.32	0.32	0.32

Table 5 Contraction numbers of the V-cycle algorithm for the Lamé problem on the unit square with $\mu = 1$ and $\lambda = 500$

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.92	0.94	0.94	0.94	0.94	0.94
(2, 2)	0.93	0.94	0.94	0.94	0.94	0.93
(4, 4)	0.88	0.90	0.90	0.90	0.90	0.90
(8, 8)	0.81	0.83	0.84	0.84	0.84	0.84
(16, 16)	0.70	0.73	0.73	0.73	0.73	0.73
(32, 32)	0.53	0.55	0.54	0.55	0.55	0.55
(64, 64)	0.31	0.32	0.32	0.32	0.32	0.32

Table 6 Contraction numbers of the W-cycle algorithm for the Lamé problem on the L-shaped domain with $\mu = 1$ and $\lambda = 500$

$(m_1, m_2)\backslash k$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
(1, 1)	0.94	0.95	0.95	0.95	0.95	0.95
(2, 2)	0.94	0.95	0.95	0.95	0.95	0.95
(4, 4)	0.91	0.92	0.92	0.92	0.92	0.92
(8, 8)	0.85	0.87	0.87	0.87	0.87	0.87
(16, 16)	0.76	0.79	0.79	0.79	0.79	0.79
(32, 32)	0.62	0.64	0.64	0.64	0.64	0.64

In addition we also list in Tables 4, 5 and 6 the contraction numbers of the multigrid W-cycle and V-cycle algorithms for the Lamé problem on both the unit square and the L-shaped domain. We take $\mu = 1$ and $\lambda = 500$ to be the Lamé constants. It is clear that the proposed multigrid iterations converge uniformly independent of the grid level on both domains and the contraction numbers decrease when we increase

Table 7 Contraction numbers of the W-cycle algorithm for the Lamé problem on the unit square with $\mu = 1$ and $k = 5$

$(m_1, m_2)\backslash\lambda$	$\lambda = 10^0$	$\lambda = 10^1$	$\lambda = 10^2$	$\lambda = 10^3$
(1, 1)	0.96	0.95	0.92	0.93
(2, 2)	0.93	0.93	0.92	0.93
(4, 4)	0.87	0.87	0.88	0.90
(8, 8)	0.76	0.79	0.81	0.83
(16, 16)	0.62	0.67	0.68	0.72
(32, 32)	0.48	0.47	0.49	0.55

the number of smoothing steps. The rate of decrease for the contraction numbers is also approaching the asymptotic rate of $(m_1 m_2)^{-1/2}$.

Finally in Table 7 we present the contraction numbers of the W-cycle algorithm for the Lamé problem on the unit square for $k = 5$ and various values of λ and $m(= m_1 = m_2)$. The robustness of the performance with respect to λ is clearly visible.

7 Concluding remarks

In this paper we develop new multigrid methods for a class of saddle point problems that include the Stokes system in fluid flow and the Lamé system in linear elasticity as special cases. The crucial ingredients are two new smoothers that take advantage of the discrete inf-sup condition and duality, which allows the convergence analysis of the W-cycle algorithm to be carried out by techniques originally invented for SPD problems.

We have established the uniform convergence of the W-cycle algorithm for a sufficiently large number of smoothing steps. But numerical results indicate that the W-cycle and V-cycle algorithms are uniformly convergent with only one smoothing step. Since our new multigrid methods can be analyzed through the connection of the saddle point problem to an equivalent SPD problem, we expect that the multiplicative convergence theory for the V-cycle algorithm with one smoothing step (cf. [10,45,46] and the references therein) can be extended to our V-cycle method, and that the decrease of the contraction number as the number of smoothing steps increases can be established by extending the additive convergence theory in [14,15].

Our approach can also be applied to other saddle point problems arising from the discretizations of boundary value problems by mixed finite element methods [31]. The key is to exploit the inf-sup condition and the multigrid theory for SPD problems associated with the relevant function spaces.

As mentioned in Remark 1.3, there are many preconditioned iterative methods for saddle point problems. For example one can apply the preconditioned MINRES [22] algorithm with $\mathbb{S}_h$ as the preconditioner. It turns out that the preconditioned MINRES is about twice as fast as our multigrid method, but it also requires about twice the amount of memory as the multigrid method, if no matrix is formed in the implementation of the two algorithms. Note that our multigrid method can be applied to certain related nonsymmetric saddle point problems, such as those arising from mixed finite element methods for the Oseen equation, with similar performance and identical memory requirement (cf. [8] for the case of convection–diffusion equation). On the other hand, if the preconditioned MINRES is replaced by the (restarted) preconditioned GMRES for such problems, then even more memory would be required and the performance would also suffer from the many matrix-vector products that must be carried out in each iteration. It would be interesting to compare the performance of our new multigrid methods with other methods for symmetric and nonsymmetric saddle point problems related to the Stokes and Lamé systems. We note that some numerical comparisons for iterative methods for the Stokes problem have been carried out in [30,35].

8 Appendix: Proof of Proposition 2.1

Since the case where $\alpha_L = 1$ is already covered by the results in [2,17], we will focus on the case where $\alpha_L < 1$. We will use C to denote a generic positive constant that only depends on Ω, μ_0, μ_1 and λ_0.

According to [1, Theorem 3.1], there exists a function $\boldsymbol{w} \in [H^{1+\alpha}(\Omega)]^d \cap [H_0^1(\Omega)]^2$ such that

$$\nabla \cdot \boldsymbol{w} = \nabla \cdot \boldsymbol{u} \quad \text{and} \quad \|\boldsymbol{w}\|_{H^{1+\alpha}(\Omega)} \le C\|\nabla \cdot \boldsymbol{u}\|_{H^{\alpha}(\Omega)}. \tag{8.1}$$

Let $z = \boldsymbol{u} - \boldsymbol{w}$ and $r = -(\mu + \lambda)(\nabla \cdot \boldsymbol{u})$. Then $(z, r) \in [H_0^1(\Omega)]^d \times L_2^0(\Omega)$ is the solution of the following Stokes problem:

$$\mu \int_\Omega \nabla z : \nabla \boldsymbol{v}\, dx - \int_\Omega (\nabla \cdot \boldsymbol{v}) r\, dx = \langle \boldsymbol{f} + \mu \Delta \boldsymbol{w}, \boldsymbol{v}\rangle \qquad \forall\, \boldsymbol{v} \in [H_0^1(\Omega)]^d, \tag{8.2a}$$

$$- \int_\Omega (\nabla \cdot z) q\, dx = 0 \qquad \forall\, q \in L_2^0(\Omega). \tag{8.2b}$$

Remark 8.1 In the derivation of (8.2) we have used the fact that an alternative weak formulation for the Lamé problem is to find $\boldsymbol{u} \in [H_0^1(\Omega)]^d$ such that

$$\mu \int_\Omega \nabla \boldsymbol{u} : \nabla \boldsymbol{v}\, dx + (\mu + \lambda) \int_\Omega (\nabla \cdot \boldsymbol{u})(\nabla \cdot \boldsymbol{v})\, dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}\, dx \qquad \forall\, \boldsymbol{v} \in [H_0^1(\Omega)]^d. \tag{8.3}$$

Since $0 \le \alpha \le \alpha_L \le \alpha_S$, the estimate (2.4) for the Stokes problem (8.2) yields

$$\|z\|_{H^{1+\alpha}(\Omega)} + \|r\|_{H^{\alpha}(\Omega)} \le C(\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} + \mu\|\Delta \boldsymbol{w}\|_{H^{-1+\alpha}(\Omega)}),$$

which together with (8.1) and the definitions of z and r implies

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + (\mu + \lambda)\|\nabla \cdot \boldsymbol{u}\|_{H^{\alpha}(\Omega)} \le C(\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} + (1 + \mu)\|\nabla \cdot \boldsymbol{u}\|_{H^{\alpha}(\Omega)}).$$

We conclude from the last estimate that

$$\|\boldsymbol{u}\|_{H^{1+\alpha}(\Omega)} + \lambda\|\nabla\cdot\boldsymbol{u}\|_{H^{\alpha}(\Omega)} \le C\|\boldsymbol{f}\|_{H^{-1+\alpha}(\Omega)} \tag{8.4}$$

provided $\lambda \ge \lambda_1$, where λ_1 is a large number depending only on Ω, μ_0 and μ_1.

The estimate (2.6) follows from (2.5) and (8.4) since $p = -\lambda(\nabla\cdot\boldsymbol{u})$.

References

1. Arnold, D.N., Scott, L.R., Vogelius, M.: Regular inversion of the divergence operator with Dirichlet boundary conditions on a polygon. Ann. Scuola Norm. Sup. Pisa Cl. Sci. **15**(4), 169–192 (1989)
2. Bacuta, C., Bramble, J.H.: Regularity estimates for solutions of the equations of linear elasticity in convex plane polygonal domains. Z. Angew. Math. Phys. **54**, 874–878 (2003)
3. Bank, R.E., Dupont, T.F.: An optimal order process for solving finite element equations. Math. Comp. **36**, 35–51 (1981)
4. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems. Acta Numerica **14**, 1–137 (2005)
5. Boffi, D., Brezzi, F., Demkowicz, L.F., Durán, R.G., Falk, R.S., Fortin, M.: Lecture Notes in Mathematics. Mixed finite elements, compatibility conditions, and applications. Springer-Verlag, Berlin (2008)
6. Braess, D., Sarazin, R.: An efficient smoother for the Stokes equation. Appl. Numer. Math. **23**, 3–19 (1996)
7. Bramble, J.H.: Multigrid Methods. Longman Scientific & Technical, Essex (1993)
8. Bramble, J.H., Kwak, D.Y., Pasciak, J.E.: Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems. SIAM J. Numer. Anal. **31**, 1746–1763 (1994)
9. Bramble, J.H., Xu, J.: Some estimates for a weighted L^2 projection. Math. Comp. **56**, 463–476 (1991)
10. Bramble, J.H., Zhang, X.: The analysis of multigrid methods. In: Ciarlet, P.G., Lions, J.L. (eds.) Handbook of Numerical Analysis, VII, pp. 173–415. North-Holland, Amsterdam (2000)
11. Brenner, S.C.: A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. SIAM J. Numer. Anal. **30**, 116–135 (1993)
12. Brenner, S.C.: A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity. Math. Comp. **63**(435–460), S1–S5 (1994)
13. Brenner, S.C.: Multigrid methods for parameter dependent problems. Math. Model. Anal. Numer. **30**, 265–297 (1996)
14. Brenner, S.C.: Convergence of the multigrid V-cycle algorithm for second order boundary value problems without full elliptic regularity. Math. Comp. **71**, 507–525 (2002)
15. Brenner, S.C.: Convergence of nonconforming V-cycle and F-cycle multigrid algorithms for second order elliptic boundary value problems. Math. Comp. **73**, 1041–1066 (2004). (electronic)
16. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer-Verlag, New York (2008)
17. Brenner, S.C., Sung, L.-Y.: Linear finite element methods for planar linear elasticity. Math. Comp. **59**, 321–338 (1992)
18. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer-Verlag, New York (1991)
19. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
20. Dauge, M.: Lecture Notes in Mathematics. Elliptic boundary value problems on corner domains, vol. 1341. Springer-Verlag, Berlin (1988)
21. Duvaut, G., Lions, J.L.: Inequalities in Mechanics and Physics. Springer-Verlag, Berlin (1976)
22. Elman, H., Silvester, D., Wathen, A.: Finite Elements and Fast Iterative Solvers. Oxford University Press, Oxford (2005)
23. Elman, H.C., Golub, G.H.: Inexact and preconditioned Uzawa algorithms for saddle point problems. SIAM J. Numer. Anal. **31**, 1645–1661 (1994)
24. Girault, V., Raviart, P.-A.: Finite Element Methods for Navier–Stokes Equations. Theory and algorithms. Springer-Verlag, Berlin (1986)
25. Grisvard, P.: Elliptic Problems in Non Smooth Domains. Pitman, Boston (1985)
26. Grisvard, P.: Singularitiés en elasticité. Arch. Ration. Mech. Anal. **107**, 157–180 (1989)

27. Grisvard, P.: Singularities in Boundary Value Problems. Masson, Paris (1992)
28. Hackbusch, W.: Multi-grid Methods and Applications. Springer-Verlag, Berlin (1985)
29. Kozlov, V.A., Maz'ya, V.G., Rossmann, J.: Spectral Problems Associated with Corner Singularities of Solutions to Elliptic Problems. AMS, Providence, USA (2001)
30. Larin, M., Reusken, A.: A comparative study of efficient iterative solvers for generalized Stokes equations. Numer. Linear Algebra Appl. **15**, 13–34 (2008)
31. Mardal, K.-A., Winther, R.: Preconditioning discretizations of systems of partial differential equations. Numer. Linear Algebra Appl. **18**, 1–40 (2011)
32. Maz'ya, V., Rossmann, J.: Elliptic Equations in Polyhedral Domains. American Mathematical Society, Providence, RI, USA (2010)
33. McCormick, S.F. (ed.): Multigrid methods. Frontiers in Applied Mathematics, vol. 3. SIAM, Philadelphia (1987)
34. Olshanskii, M.A.: Multigrid analysis for the time dependent Stokes problem. Math. Comp. **81**, 57–79 (2012)
35. Peters, J., Reichelt, V., Reusken, A.: Fast iterative solvers for discrete Stokes equations. SIAM J. Sci. Comput. **27**, 646–666 (2005)
36. Schöberl, J., Zulehner, W.: On Schwarz-type smoothers for saddle point problems. Numer. Math. **95**, 377–399 (2003)
37. Tartar, L.: An Introduction to Sobolev Spaces and Interpolation Spaces. Springer, Berlin (2007)
38. Toselli, A., Widlund, O.B.: Domain Decomposition Methods—Algorithms and Theory. Springer, New York (2005)
39. Triebel, H.: Interpolation Theory, Function Spaces, Differential Operators. North-Holland, Amsterdam (1978)
40. Trottenberg, U., Oosterlee, C., Schüller, A.: Multigrid. Academic Press, San Diego (2001)
41. Vassilevski, P.S.: Multilevel Block Factorization Preconditioners. Springer, New York (2008)
42. Verfürth, R.: A multilevel algorithm for mixed problems. SIAM J. Numer. Anal. **21**, 264–271 (1984)
43. Verfürth, R.: Multigrid algorithms for mixed problems, II. Treatment of the mini-element. SIAM J. Numer. Anal. **25**, 285–293 (1988)
44. Wittum, G.: On the convergence of multigrid methods with transforming smoothers. Numer. Math. **57**, 15–38 (1990)
45. Xu, J., Zikatanov, L.: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Am. Math. Soc **15**, 573–597 (2002)
46. Yserentant, H.: Old and new convergence proofs for multigrid methods. Acta Numerica **2**, 285–326 (1993)
47. Zulehner, W.: A class of smoothers for saddle point problems. Computing **65**, 227–246 (2000)