Computers and Mathematics with Applications 70 (2015) 2854–2866

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Multigrid methods for saddle point systems using constrained smoothers

Long Chen

Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA

ARTICLE INFO

Article history:
Received 12 December 2014
Received in revised form 29 August 2015
Accepted 20 September 2015
Available online 21 October 2015

Keywords:
Saddle point system
Multigrid methods
Mixed finite elements
Stokes equations

ABSTRACT

The constrained smoother for solving the saddle point system arising from the constrained minimization problem is a relaxation scheme such that the iteration remains in the constrained subspace. A multigrid method using constrained smoothers for saddle point systems is analyzed in this paper. Uniform convergence of two-level and W-cycle multigrid methods, with sufficient many smoothing steps and full regularity assumptions, are obtained for some stable finite element discretization of Stokes equations. For Braess–Sarazin smoother, a convergence theory using only partial regularity assumption is also developed.

1. Introduction

Due to their indefiniteness and poor spectral properties, saddle point problems are difficult to solve. Multigrid (MG) methods, one of the most efficient solvers for symmetric positive definite problems, work less efficiently for saddle point problems. In this paper, we shall design and analyze effective multigrid methods for the saddle point problems:

$$\begin{pmatrix} A & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \tag{1}$$

where A is a symmetric and positive definite (SPD) operator and B is surjective. Eq. (1) arises from mixed finite element methods discretization of partial differential equations (PDEs), notably the Stokes equations in fluid dynamics in which $A = -\Delta$, and $B = -\text{div}$.

The main difficulty of developing robust and effective multigrid methods for the saddle point system (1) is due to the constraint $Bu = 0$. Recall that the success of multigrid method relies on two ingredients: the high frequency can be damped efficiently by the smoother, and the low frequency can be well approximated by the coarse grid correction. For saddle point systems, however, both smoothing and coarse grid correction can easily violate the constraint.

We propose to use the constrained smoother which is defined as a relaxation scheme such that the iteration remains in the constrained subspace $\mathcal{K} = \ker(B)$. For Stokes equations, this means that the velocity iteration is always divergence free. To derive constrained smoothers, we introduce the operator $A_{\mathcal{K}} = Q_{\mathcal{K}} A Q_{\mathcal{K}}^T : \mathcal{K} \to \mathcal{K}$, where $Q_{\mathcal{K}} = I - B^T(BB^T)^{-1}B$ is the L^2-projection to $\mathcal{K}$, and rewrite the saddle point system (1) as a symmetric positive definite equation

$$A_{\mathcal{K}} u = Q_{\mathcal{K}} f. \tag{2}$$

We shall design smoothers for (1) based on smoothers for (2). The operator $A_{\mathcal{K}}$ is introduced for the theoretical propose and will not be formed explicitly. Namely the algorithm we derived will involve only components of the original saddle point system.

E-mail address: chenlong@math.uci.edu.

http://dx.doi.org/10.1016/j.camwa.2015.09.020
0898-1221/

We shall show that Richardson iteration for solving (2) is the Braess–Sarazin (B–S) smoother [1] for (1), and Jacobi and Gauss–Seidel iterations for (2) correspond to additive and multiplicative Schwarz smoothers considered in Schöberl [2], which is better known as Vanka smoother [3] in the context of computational fluid dynamics.

For the coarse grid correction, the difficulty is that the constrained subspaces in consecutive levels are non-nested. To overcome this non-nestedness, we propose to use a L^2-type projection $\mathbb{Q}_{\mathcal{K}}$ to bring the coarse grid correction back to $\mathcal{K}$. One generic choice is $\mathbb{Q}_{\mathcal{K}} = Q_{\mathcal{K}}$ which requires a Poisson type solver. When the pressure space consists of discontinuous elements, following Schöberl [2], we can choose a localized L^2 projection by using elements in the coarse grid as a non-overlapping domain decomposition of the underlying domain. We would like to mention that when constrained subspaces are nested, a multilevel method based on the constrained energy minimization and its convergence analysis has been developed recently in [4].

It has been numerically observed that multiplicative Schwarz smoother leads to an efficient multigrid methods for saddle point problems, however, theoretical analysis for the convergence is only available for less efficient additive versions [2,5]. One contribution of this paper is to extend the smoothing property of the additive Schwarz smoother established by Schöberl [2] to the multiplicative Schwarz smoothers.

With the smoothing property and the approximation property, we are able to prove that the two-level method and W-cycle multigrid method using constrained smoothers converge uniformly provided the full regularity assumption of Stokes equations and the assumption of sufficiently many smoothing steps.

Another contribution of this paper is to present a multigrid convergence proof without the full regularity assumption. For scalar elliptic equations, the MG theory has undergone stages of development from regularity based multigrid theory [6] to regularity free (or less) one [7–12]. Surprisingly enough the current MG theory for saddle point problems is still in the full regularity stage [13,1,14–16]. Only very recently, Brenner, Li, and Sung [17] developed new multigrid methods for Stokes equations and have proved the uniform convergence without the full regularity assumption. Our smoother and consequently the convergence analysis is different with that in [17].

We shall follow Bank and Dupont [7] to present a convergence proof using only partial regularity assumption of the Stokes equation. Consequently our analysis can be applied to more realistic problems especially for solutions with singularity. We verify the approximation and smoothing property using an operator dependent norm for the Braess–Sarazin smoother. We shall also follow Bramble, Pasciak, and Xu [18] to use the variable V-cycle multigrid as a preconditioner which can relax the assumption of sufficiently many smoothing steps.

We are aware that more effective block preconditioners for the Stokes equations are available [19,20]. The analysis here is of theoretical interest since the convergence of multigrid methods for Stokes equations with the partial regularity assumption is rare.

The rest of this paper is structured as follows. In Section 2, we present the setting of the problem including notation and different formulations of the saddle point system. In Section 3, we introduce constrained relaxation schemes and in Section 4, we present the two-level method and W-cycle multigrid and prove their uniform convergence. In Section 5, we verify the smoothing and approximation property for Vanka smoothers and in Section 6, we establish the convergence theory with partial regularity assumption when using Braess–Sarazin smoother. We refer the reader to [1,2] for numerical results that are consistent and supporting our theoretical results.

2. Problem setting

Let $\mathcal{H}$ be a Hilbert space equipped with inner product $(\cdot,\cdot)$ and $\mathcal{V} \subset \mathcal{H}$ be a closed subspace. Suppose $A : \mathcal{V} \to \mathcal{V}$ is a symmetric and positive definite (SPD) operator with respect to $(\cdot,\cdot)$, which introduces a new inner product $(u, v)_A := (Au, v) = (u, Av)$ on $\mathcal{V}$. The norm associated to $(\cdot,\cdot)$ or $(\cdot,\cdot)_A$ will be denoted by $\|\cdot\|$ or $\|\cdot\|_A$, respectively. Let $\mathcal{P}$ be another Hilbert space and let $B : \mathcal{V} \to \mathcal{P}$ be a linear operator continuous in $\|\cdot\|_A$. With a slight abuse of notation, we still denote the inner product of $\mathcal{P}$ by $(\cdot,\cdot)$. In most problems of consideration, the inner product $(\cdot,\cdot)$ for $\mathcal{H}$ is the vector L^2-inner product while for $\mathcal{P}$ it is the scalar L^2-inner product.

We are interested in solving the following saddle point system: For a given $f \in \mathcal{H}$, find $u \in \mathcal{V}$, $p \in \mathcal{P}$ such that

$$\begin{aligned} (Au, v) + (p, Bv) &= (f, v) && \text{for all } v \in \mathcal{V},\\ (Bu, q) &= 0 && \text{for all } q \in \mathcal{P}, \end{aligned}$$

which will be written in the operator form

$$\begin{pmatrix} A & B^T \\ B & O \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{3}$$

The operator matrix in (3) will be abbreviated as $\mathcal{L} = (A, B^T; B, O)$ and Eq. (3) can be written as $\mathcal{L}(u, p) = (f, 0)$. Here $(\cdot)^T$ is the adjoint with respect to the default inner product $(\cdot,\cdot)$ and a functional in the dual space $\mathcal{H}'$ is identified as an element in $\mathcal{H}$ through the Riesz map induced by $(\cdot,\cdot)$. Throughout this paper, we assume the well-posedness of (3) and focus on its efficient solvers.

We shall consider geometric multigrid methods for solving the saddle point problem (3) which arises from mixed finite element method discretizations of elliptic partial differential equations. A typical and important example is the finite element discretization of Stokes equations $-\Delta u + \nabla p = f,\ \operatorname{div} u = 0$ posed on a polygon or polyhedron domain Ω. For another important example: Darcy systems using $H(\operatorname{div})$ elements, we refer the reader to [4].

Let $\mathcal{T}_h$ be a quasi-uniform mesh of Ω with mesh size h. We consider geometric multigrid methods in this paper and thus assume $\mathcal{T}_h$ is obtained by uniform refinements from an initial mesh $\mathcal{T}_1$ of Ω, i.e., there exists a sequence of meshes $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_J = \mathcal{T}_h$. The mesh $\mathcal{T}_1$ is a shape regular mesh of Ω and $\mathcal{T}_{k+1}$ is obtained by dividing each element in $\mathcal{T}_k$ into small elements following appropriate refinement rules for different types of meshes. The mesh size $\mathcal{T}_k$ will be denoted by h_k. By the construction $h_k/h_{k+1} = 2$. We are interested in solving the system in the finest grid but coarse grids will be used to construct auxiliary problems such that only relaxation on each level is enough to produce an iterative solver convergent with a rate independent of the mesh size.

Let $\mathcal{K} = \ker(B)$ be the null space of B. The saddle point problem (3) can be reformulated to the following symmetric and positive definite (SPD) problem in $\mathcal{K}$: Find $u \in \mathcal{K}$ such that

$$(Au, v) = (f, v) \quad \text{for all } v \in \mathcal{K}. \tag{4}$$

We introduce the operator $A_{\mathcal{K}} : \mathcal{K} \to \mathcal{K}$ as $(A_{\mathcal{K}} u, v) = (Au, v)$ for all $u, v \in \mathcal{K}$ and the operator $Q_{\mathcal{K}} : \mathcal{H} \to \mathcal{K}$ as the $(\cdot,\cdot)$-projection, i.e., for a given function $f \in \mathcal{H},\ Q_{\mathcal{K}} f \in \mathcal{K}$ satisfies $(Q_{\mathcal{K}} f, v) = (f, v)$ for all $v \in \mathcal{K}$. Then the operator form of (4) is: Find $u \in \mathcal{K}$ such that

$$A_{\mathcal{K}} u = Q_{\mathcal{K}} f \quad \text{in } \mathcal{K}. \tag{5}$$

Last we define the Dirichlet energy:

$$\mathcal{E}(v) = \frac{1}{2}\|v\|_A^2 - (f, v), \quad \text{for } v \in \mathcal{V}. \tag{6}$$

Eq. (5) is Euler's equation of the following constrained minimization problem:

$$\min_{v \in \mathcal{K}} \mathcal{E}(v). \tag{7}$$

Equivalently the saddle point problem (3) is the first order equation of (7) by introducing the Lagrangian multiplier p to impose the constraint $Bv = 0$.

We shall switch our viewpoint from these three equivalent formulations: energy minimization in the constrained subspace, the SPD problem in the constrained subspace, and the saddle point system in the non-constrained space.

We use notation $a \lesssim b$ to denote there exists a positive constant C independent of the mesh size h, such that $a \le Cb$, and $a \eqsim b$ to denote $a \lesssim b \lesssim a$.

We use the standard definition of Sobolev spaces $H^s(\Omega)$ and $H_0^s(\Omega)$ with $s \ge 0$. When $s = 0,\ H^0(\Omega)$ coincides with the space of square integrable functions $L^2(\Omega)$.

3. Constrained relaxation of saddle point problems

The constrained relaxation for solving (3) is defined as a relaxation scheme such that the iteration remains in the constrained subspace. For Stokes equations, this means that the velocity iteration is always divergence free. We will always denote by u the solution to (3) and (5) and by u^k the kth iteration of u for $k = 0, 1, 2, \dots$. We chose an initial guess $u^0 \in \mathcal{K}$. A trivial example is $u^0 = 0$. Then, for a constrained relaxation, all $u^k \in \mathcal{K}$ for $k = 1, 2, \dots$.

We first explore the relation between operator $A : \mathcal{V} \to \mathcal{V}$ and $A_{\mathcal{K}} : \mathcal{K} \to \mathcal{K}$. Let $I_{\mathcal{K}} : \mathcal{K} \to \mathcal{V}$ be the natural inclusion operator. By definition $I_{\mathcal{K}} = Q_{\mathcal{K}}^T$. Then it is easy to see that $A_{\mathcal{K}} = Q_{\mathcal{K}} A I_{\mathcal{K}} = Q_{\mathcal{K}} A Q_{\mathcal{K}}^T$.

Classical iterative methods for solving $A_{\mathcal{K}} u = Q_{\mathcal{K}} f$ would require the explicit form of $A_{\mathcal{K}}$. Since $Q_{\mathcal{K}} = I - B^T(BB^T)^{-1}B$ involving an inverse operator, $A_{\mathcal{K}}$ may not be easy or efficient to form explicitly. We shall use the equivalence between (3) and (5) to design smoothers for (3) without forming $A_{\mathcal{K}}$.

Remark 3.1. For some discrete Stokes systems, it is possible to find a so-called discrete divergence free basis of $\mathcal{K}$ and each basis function is locally supported; see, e.g. [21, page 267]. Therefore a sparse representation of $A_{\mathcal{K}}$ can be obtained using this basis. The finding of such bases, however, is not easy and is problem dependent. More importantly, the condition number of the reduced SPD system is much worse than that of the saddle point system. Essentially it is a fourth order elliptic equation with an $\mathcal{O}(h^{-4})$ condition number rather than a constrained second order elliptic equation with an $\mathcal{O}(h^{-2})$ condition number. Thus even if a sparse representation of $A_{\mathcal{K}}$ is available, it is still better to solve the original saddle point system. $\square$

We first transfer the simplest iterative method, i.e., Richardson iteration for (5) to an iteration for solving (3). Recall that Richardson iteration can be written as:

$$u^{k+1} = u^k + \omega^{-1}(Q_{\mathcal{K}} f - A_{\mathcal{K}} u^k) = u^k + \omega^{-1} Q_{\mathcal{K}}(f - Au^k). \tag{8}$$

The error equation of (8) is

$$u - u^{k+1} = (I - \omega^{-1} A_{\mathcal{K}})(u - u^k) = Q_{\mathcal{K}}(I - \omega^{-1} A)(u - u^k). \tag{9}$$

In view of (9), we should chose $\omega^{-1} \in (0, 2/\rho_{A_{\mathcal{K}}})$ so that the iteration (8) converges, where ρ_M denotes the spectral radius of operator M.

From the minimization point of view, Richardson iteration (8) is related to a gradient method for solving the constrained minimization problem (7). The corresponding iteration for the saddle point system is: solve

$$\begin{pmatrix} \omega I & B^T \\ B & O \end{pmatrix} \begin{pmatrix} e_u \\ e_p \end{pmatrix} = \begin{pmatrix} f - Au^k \\ 0 \end{pmatrix}, \tag{10}$$

and then update $u^{k+1} = u^k + e_u$. One can compute the inverse $(\omega I, B; B^T, O)^{-1}$ to verify the relation (9). This is exactly the smoother developed by Braess and Sarazin [1].

Remark 3.2. In the implementation level, the identity operator I is realized by a mass matrix and the exact solve of (10) can be approximated by using an easily invertible matrix, e.g., using the diagonal of the matrix representation of operator A instead of the mass matrix. Then the inner product $(\cdot, \cdot)$ can be replaced by an equivalent mesh-dependent and weighted l^2-inner product of an Euclidean space which will not affect the analysis.

It is a coincidence that, for Richardson iteration, it is simply replacing A by ωI in the saddle point system which resembles the Richardson iteration for solving A. Other classical iterative methods such as Jacobi and Gauss–Seidel methods for the SPD formulation (5) correspond to more delicate formulations of the saddle point system. For example, when all operators are represented by matrices, using $(\mathrm{diag}(A), B^T; B, O)$ in (10) will not lead to Jacobi method for solving (5) which requires $\mathrm{diag}(A_{\mathcal{K}})$.

We shall interpret Gauss–Seidel or Jacobi iterations based on the subspace correction method [10]. Let

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_N, \quad \mathcal{V}_i \subset \mathcal{V},\ i = 1, \ldots, N,$$

be a space decomposition of $\mathcal{V}$ satisfying the condition

$$\mathcal{K} = \mathcal{K}_1 + \mathcal{K}_2 + \cdots + \mathcal{K}_N, \quad \text{with } \mathcal{K}_i = \mathcal{V}_i \cap \ker(B),\ i = 1, \ldots, N.$$

Note that the condition $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$ requires a careful choice of the space decomposition $\mathcal{V} = \sum_{i=1}^N \mathcal{V}_i$. Roughly speaking, each subspace $\mathcal{V}_i$ should be big enough to contain a basis function of $\mathcal{K}$ and each basis function of $\mathcal{K}$ should be contained in at least one $\mathcal{V}_i$.

Denote by $Q_i : \mathcal{V} \to \mathcal{K}_i$ the $(\cdot, \cdot)$-projection and $A^i_{\mathcal{K}} = Q_i A Q_i^T$ the restriction of A to the subspace $\mathcal{K}_i$. The parallel (additive) subspace correction method (PSC) is:

$$u^{k+1} = u^k + \sum_{i=1}^N (A^i_{\mathcal{K}})^{-1} Q_i (f - Au^k).$$

The successive (multiplicative) subspace correction (SSC) method is:

Let $v^0 = u^k$, for $i = 1, 2, \ldots, N$, solve $A^i_{\mathcal{K}} e_i = Q_i(f - Av^{i-1})$ and update $v^i = v^{i-1} + e_i$. The new iteration is $u^{k+1} = v^N$.

SSC and PSC differ in the update of the residual. In SSC, when solving the local problem in $\mathcal{K}_i$, the residual is updated while in PSC it is not. From the energy minimization point of view, SSC will always reduce the energy while PSC does not and usually an appropriate scaling factor is needed. Therefore SSC is more effective. On the other hand, PSC is more friendly to parallel computing.

It is well known that SSC corresponds to Gauss–Seidel type iteration and PSC is Jacobi type iteration [10].

We now transfer SSC or PSC for (5) to iterative methods for the saddle point system. Let $A_i : \mathcal{V}_i \to \mathcal{V}_i$ and $B_i : \mathcal{V}_i \to \mathcal{P}_i := \mathcal{P} \cap B(\mathcal{V}_i)$ be the restriction of A and B to the subspace $\mathcal{V}_i$ and $\mathcal{P}_i$, respectively. Given a residual r_i, the local problem $A^i_{\mathcal{K}} e_i = Q_i r_i$ in the subspace $\mathcal{K}_i$ corresponds to a small saddle point system in $\mathcal{V}_i \times \mathcal{P}_i$:

$$\begin{pmatrix} A_i & B_i^T \\ B_i & O \end{pmatrix} \begin{pmatrix} e_i \\ p_i \end{pmatrix} = \begin{pmatrix} r_i \\ 0 \end{pmatrix}. \tag{11}$$

Therefore SSC for solving $A_{\mathcal{K}} u = Q_{\mathcal{K}} f$ based on the space decomposition $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$ can be interpreted as a multiplicative Schwarz method for solving the saddle point problem based on the decomposition $\mathcal{V} = \sum_{i=1}^N \mathcal{V}_i$. Similarly PSC is the additive Schwarz method.

Denote by $R_{\mathcal{K}}$ the corresponding operator of SSC or PSC for solving (5), i.e.,

$$u^{k+1} = u^k + R_{\mathcal{K}}(Q_{\mathcal{K}} f - A_{\mathcal{K}} u^k).$$

Then the error equation is

$$u - u^{k+1} = (I - R_{\mathcal{K}} A_{\mathcal{K}})(u - u^k) := S_{\mathcal{K}}(u - u^k). \tag{12}$$

The operator $R_{\mathcal{K}}$ is introduced for the ease of analysis. In implementation, for a given residual r, the action $R_{\mathcal{K}} r$ can be realized by solving small saddle point problems (11) consecutively (SSC) or in parallel (PSC). Explicit formulation of $R_{\mathcal{K}}$ is not easy and not necessary, but possible, see e.g. [22,23].

4. Multigrid methods using constrained smoothers

Following the convention, we use subscript H to denote quantities associated to coarse spaces $\mathcal{V}_H \subset \mathcal{V}$ and $\mathcal{P}_H \subset \mathcal{P}$ which are usually constructed on a coarse grid with grid size $H = 2h$. We denote by $I_H : \mathcal{V}_H \to \mathcal{V}$ the natural inclusion which is usually skipped if no confusion arises. The adjoint I_H^T is the restriction (of residual) from $\mathcal{V} \to \mathcal{V}_H$. One difficulty of applying standard multigrid methods to saddle point systems is the non-nestedness of the constrained subspaces, i.e., in general $\mathcal{K}_H \not\subset \mathcal{K}$ although $\mathcal{V}_H \subset \mathcal{V}$. To overcome it, we introduce another operator $\mathbb{Q}_{\mathcal{K}} : \mathcal{V} \to \mathcal{K}$ which brings a function in $\mathcal{V}$ back to the constrained space $\mathcal{K}$ and require that $\mathbb{Q}_{\mathcal{K}}$ restricted to $\mathcal{K}$ is the identity operator. One generic choice is $\mathbb{Q}_{\mathcal{K}} = Q_{\mathcal{K}}$. The restriction of $\mathbb{Q}_{\mathcal{K}}$ to $\mathcal{K}_H$ can be thought of as a prolongation operator between non-nested spaces $\mathcal{K}_H \to \mathcal{K}$.

4.1. Two-level method and its convergence

A two-level method using a constrained smoother $R_{\mathcal{K}}$ is presented below and the multigrid V-cycle or W-cycle multigrid can be obtained by recursion.

$u^{k+1} = \mathrm{TM}\,(u^k, f)$.

Set $v_1 = u^k$.

(1) Pre-smoothing. For $i = 1, \ldots, m$, $v_{i+1} = v_i + R_{\mathcal{K}} Q_{\mathcal{K}}(f - Av_i)$.
(2) Coarse grid correction: $e_H = A_{\mathcal{K}_H}^{-1} Q_{\mathcal{K}_H} I_H^T (f - Av_{m+1})$.
(3) Prolongate the correction back to the kernel space: $e = \mathbb{Q}_{\mathcal{K}} I_H e_H$.
(4) Update the approximation: $v_{m+2} = v_{m+1} + e$.
(5) Post-smoothing. For $i = m+2, \ldots, 2m+1$, $v_{i+1} = v_i + R_{\mathcal{K}} Q_{\mathcal{K}}(f - Av_i)$.

Set $u^{k+1} = v_{2m+2}$.

In the pre- and post-smoothing, the projection of the residual $Q_{\mathcal{K}} r$, with $r = f - Av_i$, is not computed explicitly when using the constrained smoother. Since in the fine space the subspace $\mathcal{K}_i \subset \mathcal{K}$, the action $(Q_{\mathcal{K}} Ar, v_i) = (Ar, v_i)$ for $v_i \in \mathcal{K}_i \subset \mathcal{K}$ can be computed without computing $Q_{\mathcal{K}} Ar$.

In the coarse grid problem, again, evaluation of the projection $Q_{\mathcal{K}_H}$ is not needed either. The correction $e_H = A_{\mathcal{K}_H}^{-1} Q_{\mathcal{K}_H}(I_H^T r)$ will be obtained by solving the following saddle point problem in the coarse space:

$$\begin{pmatrix} A_H & B_H^T \\ B_H & O \end{pmatrix} \begin{pmatrix} e_H \\ q_H \end{pmatrix} = \begin{pmatrix} I_H^T r \\ 0 \end{pmatrix}. \tag{13}$$

After we obtained a correction in the coarse grid, say $e_H \in \mathcal{K}_H$, since $\mathcal{K}_H \not\subset \mathcal{K}$, the direct update using e_H will be out of the subspace $\mathcal{K}$. Thus in step (3), we do need to compute the projection $\mathbb{Q}_{\mathcal{K}} e_H$. A generic choice is $\mathbb{Q}_{\mathcal{K}} = Q_{\mathcal{K}}$ which requires a Poisson-type solver in the case of Stokes equations. When the space of Lagrange multiplier is discontinuous, we can construct such a $\mathbb{Q}_{\mathcal{K}}$ by solving local problems. Specific examples will be given in the next section.

If we introduce the operator $P_H : \mathcal{V} \to \mathcal{V}_H$ as $P_H = A_{\mathcal{K}_H}^{-1} Q_{\mathcal{K}_H} I_H^T A$, then the error operator of the two-level method (TM) is

$$S_{\mathcal{K}}^m (I - \mathbb{Q}_{\mathcal{K}} P_H) S_{\mathcal{K}}^m = S_{\mathcal{K}}^m \mathbb{Q}_{\mathcal{K}} (I - P_H) S_{\mathcal{K}}^m. \tag{14}$$

Here we use the property $\mathbb{Q}_{\mathcal{K}} v = v$ when $v \in \mathcal{K}$.

We shall prove the uniform convergence of this two-level method from which convergence of the W-cycle follows from the standard recursive argument. As usual, we present the following assumptions on the smoother and coarse grid correction.

(S) Smoothing property. There exists a function $\eta(m)$ with property $\lim_{m\to\infty} \eta(m) = 0$ such that

$$\|S_{\mathcal{K}}^m v\|_A \lesssim \eta(m) \rho_{A_{\mathcal{K}}}^{1/2} \|v\|, \quad \text{for all } v \in \mathcal{K}.$$

(R) The smoother $R_{\mathcal{K}}$ is symmetric, non-singular, and satisfies

$$(R_{\mathcal{K}}^{-1} v, v) \ge (A_{\mathcal{K}} v, v), \quad \text{for all } v \in \mathcal{K}.$$

(A) Approximation property of the coarse grid correction:

$$\|(I - P_H) v\| \lesssim \rho_{A_{\mathcal{K}}}^{-1/2} \|v\|_A \quad \text{for all } v \in \mathcal{V}.$$

Due to the non-nestedness of constrained subspaces, we need one more assumption of the prolongation operator $\mathbb{Q}_{\mathcal{K}}$.

(Q) The operator $\mathbb{Q}_{\mathcal{K}} : \mathcal{V} \to \mathcal{K}$ is stable in $\|\cdot\|$-norm and preserves $\mathcal{K}$. Namely $\mathbb{Q}_{\mathcal{K}}|_{\mathcal{K}}$ is identity and

$$\|\mathbb{Q}_{\mathcal{K}} v\| \lesssim \|v\|, \quad \text{for all } v \in \mathcal{V}.$$

Theorem 4.1. *Assume that*

- *the smoother $R_{\mathcal{K}}$ satisfies the assumptions (S) and (R).*
- *the coarse grid correction P_H satisfies the approximation property (A).*
- *the operator $\mathbb{Q}_{\mathcal{K}}$ satisfies the assumption (Q).*

Then the two-level method converges in A-norm with sufficiently many smoothing steps. More precisely, there exists a constant C independent of the size of the problem such that

$$\|u - u^{k+1}\|_A \le C\eta(m) \|u - u^k\|_A.$$

Proof. The assumption (R) implies $\sigma(R_{\mathcal{K}}A_{\mathcal{K}}) \in (0, 1]$. Then $S_{\mathcal{K}} = I - R_{\mathcal{K}}A_{\mathcal{K}}$ is a contraction in A-norm. Recall that the error equation is

$$u - u^{k+1} = S_{\mathcal{K}}^m \mathbb{Q}_{\mathcal{K}}(I - P_H)S_{\mathcal{K}}^m(u - u^k).$$

We then use the assumptions to estimate the A-norm of the error operator. For any $v \in \mathcal{K}$,

$$\begin{aligned}\|S_{\mathcal{K}}^m \mathbb{Q}_{\mathcal{K}}(I - P_H)S_{\mathcal{K}}^m v\|_A &\lesssim \eta(m)\rho_{A_{\mathcal{K}}}^{1/2}\|\mathbb{Q}_{\mathcal{K}}(I - P_H)S_{\mathcal{K}}^m v\| \\ &\le \eta(m)\rho_{A_{\mathcal{K}}}^{1/2}\|(I - P_H)S_{\mathcal{K}}^m v\| \lesssim \eta(m)\|S_{\mathcal{K}}^m v\|_A \lesssim \eta(m)\|v\|_A.\end{aligned}$$

The desired inequality then follows. □

4.2. W-cycle method and convergences

We present the standard recursive formulation of the W-cycle using TM. We introduce one more index l for levels. The coarsest level is $l = 1$ and the finest one is $l = J$.

$u^{k+1} =$ W-cycle (u^k, f, l)

If $l == 1$, then
  $u^{k+1} = u^k + A_{\mathcal{K}}^{-1}Q_{\mathcal{K}}(f - u^k)$; return;

else

Set $v_1 = u^k$.

(1) Pre-smoothing. For $i = 1, \dots, m$, $v_{i+1} = v_i + R_{\mathcal{K}}Q_{\mathcal{K}}(f - Av_i)$.
(2) Coarse grid correction: let $r_H = I_H^T(f - Av_{m+1})$
  (a) $\tilde{e}_H =$ W-cycle$(0, r_H, l - 1)$;
  (b) $e_H =$ W-cycle$(\tilde{e}_H, r_H, l - 1)$.
(3) Prolongate the correction back to the kernel space: $e = \mathbb{Q}_{\mathcal{K}}I_H e_H$.
(4) Update the approximation: $v_{m+2} = v_{m+1} + e$.
(5) Post-smoothing. For $i = m + 2, \dots, 2m + 1$, $v_{i+1} = v_i + R_{\mathcal{K}}Q_{\mathcal{K}}(f - Av_i)$.

Set $u^{k+1} = v_{2m+2}$.

Use the standard recursive argument, we can prove the uniform convergence of W-cycle.

Theorem 4.2. *Assume that in each level*

- *the smoother $R_{\mathcal{K}}$ satisfies the assumptions (S) and (R).*
- *the coarse grid correction P_H satisfies the approximation property (A).*
- *the operator $\mathbb{Q}_{\mathcal{K}}$ satisfies the assumption (Q).*

Then the W-cycle multigrid converges in A-norm with sufficient many smoothing steps. More precisely, there exists a constant C independent of the size of the problem such that, when m is large enough,

$$\|u - u^{k+1}\|_A \le C\eta(m)\|u - u^k\|_A.$$

5. Application to Stokes equations

In this section, we apply our approach to design and analyze multigrid methods for saddle point systems arising from finite element discretization of Stokes equations. Here we focus on stationary Stokes problems.

Consider Stokes equations with Dirichlet boundary condition posed on a polygon/polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$,

$$-\Delta u + \nabla p = f, \qquad \mathrm{div} u = 0 \quad \text{in } \Omega, \qquad u = 0, \quad \text{on } \partial\Omega. \tag{15}$$

Let $L_0^2(\Omega) := \{q \in L^2(\Omega), \int_\Omega q = 0\}$ endowed with L^2-norm $\|\cdot\|$ and L^2-inner product $(\cdot,\cdot)$, and $(H_0^1(\Omega))^d := \{u \in (L^2(\Omega))^d, \nabla u \in (L^2(\Omega))^{d\times d}, u|_{\partial\Omega} = 0\}$ with norm $|\cdot|_1 := \|\nabla(\cdot)\|$. The weak formulation of (15) is: find $u \in (H_0^1(\Omega))^d$, $p \in L_0^2(\Omega)$ such that

$$\begin{aligned}(\nabla u, \nabla v) - (p, \mathrm{div} v) &= (f, v) &&\text{for all } v \in (H_0^1(\Omega))^d,\\ -(\mathrm{div} u, q) &= 0 &&\text{for all } q \in L_0^2(\Omega).\end{aligned}$$

Given a quasi-uniform mesh $\mathcal{T}_h$ with mesh size h, we consider inf–sup stable finite element spaces. The setting is: spaces $(\mathcal{V}, \mathcal{P}) = (\mathcal{V}_h, \mathcal{P}_h)$ are stable finite element pair based on $\mathcal{T}_h$. Space $\mathcal{H} = (L^2(\Omega))^d$ with the L^2 inner-product $(\cdot,\cdot)$. The A inner-product is $(Au, v) = (\nabla u, \nabla v)$. The operator $B = -\mathrm{div}$ and $A = -\Delta$. We shall use subscript h or H, respectively, to indicate operators restricted to spaces on $\mathcal{T}_h$ or $\mathcal{T}_H$. The coarse grid size $H = 2h$.

It is well known that $\rho_{A_h}, \rho_{A_{\mathcal{K}_h}} = \mathcal{O}(h^{-2})$.

5.1. Approximation property

In this subsection, we denote the exact solution of Stokes equations by (u, p) and the finite element approximation based on triangulation $\mathcal{T}_h$ by (u_h, p_h). Notice that for all existing stable finite element pairs, we have first order convergence of u and p, i.e.,

$$\|u - u_h\|_1 + \|p - p_h\| \lesssim h\,(\|u\|_2 + \|p\|_1)\,. \tag{16}$$

To verify the approximation property in L^2-norm, we need the following full regularity assumption.

(Reg) Let $u \in (H_0^1(\Omega))^d$, $p \in L_0^2(\Omega)$ be the weak solution of Stokes equations (15) with a given data $f \in L^2(\Omega)$. Then $u \in (H^2(\Omega) \cap H_0^1(\Omega))^d$ and $p \in H^1(\Omega)$ and

$$\|u\|_2 + \|p\|_1 \lesssim \|f\|.$$

Using the standard duality argument and the above full regularity assumption, we can get the following approximation result in L^2-norm:

$$\|u - u_h\| \lesssim h^2\|f\|. \tag{17}$$

We then verify the approximation property (A). Note that we are solving the discrete Stokes equation on $\mathcal{T}_h$ and the coarse spaces are based on $\mathcal{T}_H$. The approximation property is given by the error estimate between two consecutive meshes $\mathcal{T}_h$ and $\mathcal{T}_H$.

Theorem 5.1. *Assume (Reg) holds for Stokes equations. Let $P_H : \mathcal{V}_h \to \mathcal{V}_H$ be defined as $P_H = A_{\mathcal{K}_H}^{-1} Q_{\mathcal{K}_H} I_H^T A_h$. Then the following approximation property holds:*

$$\|v_h - P_H v_h\| \lesssim h\|v_h\|_A \quad \text{for all } v_h \in \mathcal{V}_h.$$

Proof. Let $f = A_h v_h$ and (v, p) be the solution of Stokes equations with data f. Then v_h and $v_H = P_H v_h$ are the Galerkin approximation of v in $\mathcal{V}_h$ and $\mathcal{V}_H$, respectively. By the triangle inequality, the L^2-error estimate (17), the inverse inequality, and the fact $H/h \le C$, we have

$$\|v_h - P_H v_h\| \le \|v - v_h\| + \|v - v_H\| \lesssim H^2\|Av_h\| \lesssim H\|v_h\|_A \lesssim h\|v_h\|_A.$$

This completes the proof. □

5.2. Constrained smoothers

To define a constrained smoother, it suffices to give a space decomposition $\mathcal{V} = \sum_{i=1}^{N_h} \mathcal{V}_i$ such that $\mathcal{K} = \sum_{i=1}^{N_h} \mathcal{K}_i$ with $\mathcal{K}_i = \mathcal{V}_i \cap \ker(B)$. The key of an effective constrained smoother is to identify the support of the local spaces $\mathcal{V}_i$. On the one side, the size of the local problem should be small such that local problems are efficient to solve. On the other side, the space $\mathcal{V}_i$ should be big enough to enclose a basis function of $\mathcal{K}$. Note that we do not need to figure out a discrete divergence-free basis but only need to know the support of a basis function. In the following, we will give the sub-domain ω for specific examples and the local problem is solving Stokes equations with Dirichlet boundary condition on $\partial\omega$. We refer the reader to Sarin [24] for algebraic ways to identify patches of divergence free basis which might be helpful on designing algebraic multigrid methods for Stokes problems.

A rule of thumb is that, when the pressure space is discontinuous, chose the vertex patch of the triangulation. More specifically, when the pressure is piecewise constant, the velocity is quadratic Lagrange element or Crouzeix–Raviart (CR) non-conforming linear finite element space [25], the support of two type of discrete divergence free functions (vertex-type and edge-type) will be contained in the vertex patch; see [26,27]. On rectangular grids, for the $Q_2 - P_1$ (bi-quadratic velocity and discontinuous linear pressure) pair, the vertex patch also contains the support of four types of discrete divergence free functions; see [21, page 271].

For $Q_2 - Q_1$ (continuous bi-quadratic velocity and continuous bi-linear pressure), we can still use the vertex patch which consists of a 2×2 sub-mesh and contains the support of discrete divergence free basis functions [28]. For lowest order Taylor–Hood $P_2 - P_1$ elements, i.e., the velocity is continuous quadratic element and the pressure is continuous linear element, in addition to the vertex patch, we need to solve additional local problems in a triangle patch which is defined as the union of triangles sharing edges with a given triangle. The discrete divergence free basis will be contained either in the vertex patch or the triangle patch; see [29,30].

5.3. Prolongation operator

We discuss several choices of $\mathbb{Q}_{\mathcal{K}}$. One universal choice will be the L^2 projection to $\mathcal{K}$. Namely we chose $\mathbb{Q}_{\mathcal{K}} = Q_{\mathcal{K}}$ which requires a Poisson solver. As remarked in [1], this is practical since a lot of efficient Poisson solvers are available and the Poisson equations need only be solved with low accuracy.

When the pressure space is discontinuous, we can chose local L^2 (or H^1)-projections. For each triangle $T \in \mathcal{T}_H$, we consider the L^2-projection restricted to $\mathcal{K}_T = \{v \in \mathcal{K}, \operatorname{supp} v \subseteq \bar{T}\}$. A similar local H^1-projection was proposed by Schöberl in [2].

The stability of $\mathbb{Q}_{\mathcal{K}}$ in L^2-norm is trivial. When v is already in $\mathcal{K}$, the solution to the local problems is itself and thus this operator preserves functions $v \in \mathcal{K}$.

5.4. Smoothing property of additive Schwarz smoothers

A smoothing property for additive Schwarz smoothers has been established by Schöberl [2]. Here we review his approach briefly and will extend to the multiplicative case in the next subsection.

Assume the space decomposition $\mathcal{V} = \sum_{i=1}^N \mathcal{V}_i$ satisfying $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$ with $\mathcal{K}_i = \mathcal{V}_i \cap \mathcal{K}$. The additive Schwarz smoother $R_{\mathcal{K}}$ based on this space decomposition can be written as $R_{\mathcal{K}} = \omega D_{\mathcal{K}} := \omega \sum_{i=1}^N A_{\mathcal{K}_i}^{-1} Q_{\mathcal{K}_i}$, where the parameter ω is chosen approximately such that (*R*) is satisfied. It is well known that, e.g., [11]

$$(D_{\mathcal{K}}^{-1} v, v) =: \|v\|_{D_{\mathcal{K}}^{-1}}^2 = \inf_{\sum_{i=1}^N v_i = v, v_i \in \mathcal{K}_i} \sum_{i=1}^N \|v_i\|_A^2. \tag{18}$$

In [2], Schöberl has proved the following inequality:

$$\|v\|_{[D_{\mathcal{K}}^{-1}, A]_{1/2}} \le ch^{-1}\|v\|, \tag{19}$$

where $\|\cdot\|_{[D_{\mathcal{K}}^{-1}, A]_{1/2}}$ is the interpolation norm between $\|\cdot\|_{D_{\mathcal{K}}^{-1}}$ and $\|\cdot\|_A$ norm with parameter $1/2$.

The smoothing property

$$\|S_{\mathcal{K}}^m v\|_A \lesssim m^{-1/2}\|v\|_{D_{\mathcal{K}}^{-1}} \tag{20}$$

is well known [6] and by our choice of the relaxation parameter ω, the operator $S_{\mathcal{K}}$ is a contraction in A-norm, i.e.,

$$\|S_{\mathcal{K}}^m v\|_A \le \|v\|_A. \tag{21}$$

Then interpolation gives the smoothing property

$$\|S_{\mathcal{K}}^m v\|_A \lesssim m^{-1/4}\|v\|_{[D_{\mathcal{K}}^{-1}, A]_{1/2}} \lesssim m^{-1/4} h^{-1}\|v\|.$$

5.5. Smoothing property of multiplicative Schwarz smoothers

In this subsection, we verify the symmetric multiplicative Schwarz smoother (or Vanka smoother) satisfies assumptions (*R*) and (*S*).

Let $R_{\mathcal{K}}$ be the operator of SSC based on the space decomposition $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$. The symmetrized Vanka smoother $\overline{R}_{\mathcal{K}}$ is an operator defined by the relation

$$I - \overline{R}_{\mathcal{K}} A_{\mathcal{K}} = (I - R_{\mathcal{K}}^T A_{\mathcal{K}})(I - R_{\mathcal{K}} A_{\mathcal{K}}). \tag{22}$$

Note that the relaxation $R_{\mathcal{K}}^T$ is realized by applying multiplicative Schwarz method in the reversed ordering, i.e., if $R_{\mathcal{K}}$ is solving local problems from $i = 1, 2, \ldots, N$, then $R_{\mathcal{K}}^T$ is from $i = N,\ N-1, \ldots, 1$.

From the energy minimization point of view, the energy is strictly decreasing for the multiplicative Schwarz method which implies the contraction of the operator $\overline{S}_{\mathcal{K}} = I - \overline{R}_{\mathcal{K}}$ and therefore (*R*) holds for $\overline{R}_{\mathcal{K}}$. Indeed any symmetrized scheme defined by the relation (22) will satisfy (*R*). A short proof is as follows. Since $A_{\mathcal{K}}$ is SPD, $(I - R_{\mathcal{K}}^T A_{\mathcal{K}})(I - R_{\mathcal{K}} A_{\mathcal{K}}) = (I - R_{\mathcal{K}} A_{\mathcal{K}})^*(I - R_{\mathcal{K}} A_{\mathcal{K}})$ is SPD, where the adjoint $(\cdot)^*$ is with respect to $(\cdot,\cdot)_{A_{\mathcal{K}}}$. Therefore $\lambda_{\min}(I - \overline{R}_{\mathcal{K}} A_{\mathcal{K}}) \ge 0$ which implies $\rho(\overline{R}_{\mathcal{K}} A_{\mathcal{K}}) \le 1$. Obviously $\overline{R}_{\mathcal{K}}$ is SPD. So this verifies that $\overline{R}_{\mathcal{K}}$ satisfies (*R*).

Let $P_i : \mathcal{K} \to \mathcal{K}_i$ be the projection in $(\cdot,\cdot)_A$ inner product, i.e., $P_i v \in \mathcal{K}_i$ satisfies $(P_i v, w_i)_A = (v, w_i)_A$ for all $w_i \in \mathcal{K}_i$. We have the following characterization of the norm introduced by $\overline{R}_{\mathcal{K}}$: for all $v \in \mathcal{K}$

$$(\overline{R}_{\mathcal{K}}^{-1} v, v) = \|v\|_A^2 + \inf_{\sum_{i=0}^J v_i = v, v_i \in \mathcal{K}_i} \sum_{i=1}^N \left\| P_i \sum_{j=i+1}^J v_j \right\|_A^2. \tag{23}$$

The identity (23) can be found in [11]. A simple proof of (23) is given in [23].

Lemma 5.2. *Assume that the space decomposition $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$ is finite overlapping, i.e., the cardinality of $n(i) = \{j \in [1, N] \mid \mathcal{K}_j \cap \mathcal{K}_i \neq \varnothing\}$ is uniformly bounded for all $i = 1, 2, \ldots, N$, and the inequality (19) holds. Then $\overline{R}_{\mathcal{K}}$ satisfies the smoothing property (S).*

Proof. By the finite overlapping property, we have

$$\sum_{i=1}^N \left\| P_i \sum_{j=i+1}^J v_j \right\|_A^2 \lesssim \sum_{i=0}^N \sum_{j\in n(i)} \|v_j\|_A^2 \lesssim \sum_{i=0}^N \|v_i\|_A^2, \quad \text{and} \quad \|v\|_A^2 \lesssim \sum_{i=0}^N \|v_i\|_A^2.$$

Therefore we have, for symmetric multiplicative Schwarz smoother $\overline{R}_{\mathcal{K}}$,

$$(\overline{R}_{\mathcal{K}}^{-1} v, v) \lesssim (D_{\mathcal{K}}^{-1} v, v). \tag{24}$$

The smoothing property

$$\|S_{\mathcal{K}}^m v\|_A \lesssim m^{-1/2} \|v\|_{\overline{R}_{\mathcal{K}}^{-1}} \lesssim m^{-1/2} \|v\|_{D_{\mathcal{K}}^{-1}}, \tag{25}$$

is proved as before. Since $\overline{R}_{\mathcal{K}}$ satisfies (R), the operator $S_{\mathcal{K}}$ is a contraction in A-norm:

$$\|S_{\mathcal{K}}^m v\|_A \le \|v\|_A. \tag{26}$$

Then interpolation gives the smoothing property

$$\|S_{\mathcal{K}}^m v\|_A \lesssim m^{-1/4} \|v\|_{[D_{\mathcal{K}}^{-1}, A]_{1/2}} \lesssim m^{-1/4} h^{-1} \|v\|. \quad \square$$

The finite overlapping property is usually true for finite element methods. The inequality (19) has been established in [2] for $P_2 - P_0$ pair (continuous and piecewise quadratic element for the velocity and piecewise constant element for the pressure) and can be proved similarly for other stable pairs with discontinuous pressure spaces. Together with the approximation property proved in Theorem 5.1, we thus have proved the uniform convergence of W-cycle multigrid methods for solving Stokes equations with multiplicative Schwarz smoothers. Note that in [2], it shows numerically that the multiplicative Schwarz smoothers is more effective while the theoretical analysis in [2] is restricted to the additive case only.

6. Convergence theory with partial regularity assumption

Results obtained in the previous sections require the full regularity assumption and sufficiently many smoothing steps. In some scenario, e.g., the domain Ω is concave with a reenter corner, the full regularity assumption is violated and only partial regularity assumption holds. We shall follow Bank and Dupont [7] to prove the multigrid convergence with partial regularity assumption for the B–S smoother. The key is to verify the approximation and the smoothing property using an operator dependent norm. We shall also follow Bramble, Pasciak, and Xu [18] to use the variable V-cycle multigrid as a preconditioner which can relax the assumption of sufficiently many smoothing steps.

6.1. Fractional norm and stability

We will use a fractional norm defined by the SPD operator $A_{\mathcal{K}} : \mathcal{K} \to \mathcal{K}$. For $s \in [0, 2]$ and $v \in \mathcal{K}$, we define

$$|||v|||_s := \|A_{\mathcal{K}}^{s/2} v\|.$$

Obviously $|||v|||_0 = \|v\|$ and $|||v|||_1 = \|v\|_A$ and thus by interpolation $|||v|||_s \eqsim \|v\|_s$ for all $v \in \mathcal{K}$ and $s \in [0, 1]$.

Lemma 6.1. *The L^2 projection $Q_{\mathcal{K}}$ is stable in $|||\cdot|||_s$ for all $s \in [0, 1]$.*

Proof. By definition $|||Q_{\mathcal{K}}|||_0 = \|Q_{\mathcal{K}}\| = 1$. By Lemma 1 in [14], we have the stability of $Q_{\mathcal{K}}$ in A-norm, i.e., $|||Q_{\mathcal{K}}|||_1 = \|Q_{\mathcal{K}}\|_A \le C$. Interpolation then leads to the result. $\square$

6.2. Two-level method and its convergence

We present the two-level method using a constrained smoother $R_{\mathcal{K}}$ below.

$u^{k+1} = \text{TM-R}\,(u^k, f)$

Set $v_1 = u^k$.

(1) Pre-smoothing. For $i = 1, \ldots, m$, $v_{i+1} = v_i + R_{\mathcal{K}} Q_{\mathcal{K}} (f - A v_i)$.
(2) Project the residual to the kernel space: $r = Q_{\mathcal{K}} (f - A v_{m+1})$.
(3) Coarse grid correction: $e_H = A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T I_H^T r$.
(4) Prolongate the correction back to the kernel space: $e = Q_{\mathcal{K}} I_H I_{\mathcal{K}_H} e_H$.
(5) Update the approximation: $v_{m+2} = v_{m+1} + e$.
(6) Post-smoothing. For $i = m+2, \ldots, 2m+1$, $v_{i+1} = v_i + R_{\mathcal{K}} Q_{\mathcal{K}} (f - A v_i)$.

Set $u^{k+1} = v_{2m+2}$.

Comparing with the two-level method TM, the difference of TM-R is that: when transfer the residual to the coarse grid, we add one L^2-projection $Q_{\mathcal{K}}$ before applying the restriction from $\mathcal{V}$ to $\mathcal{V}_H$ and $\mathbb{Q}_{\mathcal{K}} = Q_{\mathcal{K}}$ in the prolongation step. The use of $Q_{\mathcal{K}}$ will increase the computational cost a little bit but the benefit is that now the error operator of TM-R is symmetric in the $(\cdot, \cdot)_A$ inner-product.

Lemma 6.2. *Let $u^{k+1} = TM\text{-}R\,(u^k, f)$. Define E by the relation $u - u^{k+1} = E(u - u^k)$. Then E is symmetric in the $(\cdot,\cdot)_A$ inner-product.*

Proof. We write out the error operator first

$$E = S_{\mathcal{K}}^m (I - Q_{\mathcal{K}} I_H I_{\mathcal{K}_H} A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T I_H^T Q_{\mathcal{K}}^T A_{\mathcal{K}}) S_{\mathcal{K}}^m, \tag{27}$$

and explain step by step. The extra projection in step (2) is to realize the action $A_{\mathcal{K}}(\cdot) = Q_{\mathcal{K}} A(\cdot)$. The restriction $I_H^T : \mathcal{V} \to \mathcal{V}_H$ is applied to elements in $\mathcal{V}$. So we need the natural inclusion $I_{\mathcal{K}} = Q_{\mathcal{K}}^T : \mathcal{K} \to \mathcal{V}$ after the action $A_{\mathcal{K}}$. The action of operator $A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T$ is realized by solving the saddle point system in the coarse grid. After we obtain $e_H \in \mathcal{K}_H$, we embed it to $\mathcal{V}_H$ through $I_{\mathcal{K}_H} : \mathcal{K}_H \to \mathcal{V}_H$ and then to $\mathcal{V}$ through $I_H : \mathcal{V}_H \to \mathcal{V}$. The projection $Q_{\mathcal{K}}$ from $\mathcal{V}$ to $\mathcal{K}$ is applied to bring the iteration back to $\mathcal{K}$.

Since $Q_{\mathcal{K}} I_H I_{\mathcal{K}_H} A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T I_H^T Q_{\mathcal{K}}^T$, $S_{\mathcal{K}}$, and $A_{\mathcal{K}}$ are symmetric operators from $\mathcal{K} \to \mathcal{K}$, we conclude $E : \mathcal{K} \to \mathcal{K}$ is symmetric in the A inner-product. □

To prove the uniform convergence of TM-R without the full regularity assumption, we modify the smoothing property to:

(S_α) Smoothing property. There exist a constant $\alpha \in (0, 1]$ and a constant c_s such that

$$\left\| S_{\mathcal{K}}^m v \right\|_{1+\alpha} \le \left(\frac{c_s}{2m}\right)^{\alpha/2} \rho_{A_{\mathcal{K}}}^{\alpha/2} \|v\|_A, \quad \text{for all } v \in \mathcal{K}. \tag{28}$$

We then define a slightly different coarse grid correction operator $P_H : \mathcal{K} \to \mathcal{K}_H \subset \mathcal{V}$ as $P_H = I_H I_{\mathcal{K}_H} A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T I_H^T Q_{\mathcal{K}}^T A_{\mathcal{K}}$ and formulate the approximation property using the fractional norm.

(A_α) Approximation property. There exist a constant $\alpha \in (0, 1]$ and a constant c_a such that

$$\|(I - P_H)v\|_{1-\alpha} \le c_a \rho_{A_{\mathcal{K}}}^{-\alpha/2} \|v\|_A, \quad \text{for all } v \in \mathcal{K}.$$

Theorem 6.3. *Assume that*

- *the smoother $R_{\mathcal{K}}$ satisfies the assumptions (R) and (S_α).*
- *the coarse grid correction P_H satisfies the approximation property (A_α).*

Then the two-level algorithms TW-R converges in the energy norm when the smoothing step m is sufficiently large. More precisely, there exists a constant $C = C(c_s, c_a, \alpha)$ such that

$$\|u - u^{k+1}\|_A \le \frac{C}{(2m)^{\alpha/2}} \|u - u^k\|_A.$$

Proof. Since the error operator $E = S_{\mathcal{K}}^m (I - Q_{\mathcal{K}} P_H) S_{\mathcal{K}}^m : \mathcal{K} \to \mathcal{K}$ is symmetric with respect to $(\cdot,\cdot)_A$, we can estimate $\|E\|_A$ by proving $(Ev, v)_A \lesssim (v, v)_A$ for all $v \in \mathcal{K}$ as follows:

$$\begin{aligned}
(Ev, v)_A &= ((I - Q_{\mathcal{K}} P_H) S_{\mathcal{K}}^m v, S_{\mathcal{K}}^m v)_A \\
&\le \left\| (I - Q_{\mathcal{K}} P_H) S_{\mathcal{K}}^m v \right\|_{1-\alpha} \left\| S_{\mathcal{K}}^m v \right\|_{1+\alpha} \\
&\le \left\| (I - P_H) S_{\mathcal{K}}^m v \right\|_{1-\alpha} \left(\frac{c_s}{2m}\right)^{\alpha/2} \rho_{A_{\mathcal{K}}}^{\alpha/2} \|v\|_A \\
&\le c_a \left(\frac{c_s}{2m}\right)^{\alpha/2} \|v\|_A^2.
\end{aligned}$$

In the third step, we write $I - Q_{\mathcal{K}} P_H = Q_{\mathcal{K}}(I - P_H)$ and use the stability of $Q_{\mathcal{K}}$ in $\|\cdot\|_{1-\alpha}$ norm; see Lemma 6.1. □

Again using the standard recursive argument, we can obtain the uniform convergence of W-cycle when smoothing steps are sufficiently large. The formulation of algorithm and results are similar and thus skipped here.

6.3. Smoothing property

We shall derive the smoothing property (S_α) from the following smoothing property (S_ρ) which is easier to verify.

(S_ρ) Smoothing property of high frequency. There exists a constant c_s such that

$$(R_{\mathcal{K}}^{-1} v, v) \le c_s \rho_{A_{\mathcal{K}}} (v, v) \quad \text{for all } v \in \mathcal{K}. \tag{29}$$

We use the notation of comparing symmetric operators: for two SPD operators M and O, we write $M \le O$ if $(Mv, v) \le (Ov, v)$ for all v in $\mathcal{K}$ or $\mathcal{V}$. We can write (29) as $R_{\mathcal{K}}^{-1} \le c_s \rho_{A_{\mathcal{K}}} I$ which is equivalent to $I \le c_s \rho_{A_{\mathcal{K}}} R_{\mathcal{K}}$. Multiplying $A_{\mathcal{K}}$ from left and right, we get another form of (S_ρ)

$$A_{\mathcal{K}}^2 \le c_s \rho_{A_{\mathcal{K}}} A_{\mathcal{K}} R_{\mathcal{K}} A_{\mathcal{K}} = c_s \rho_{A_{\mathcal{K}}} A_{\mathcal{K}} (I - S_{\mathcal{K}}),$$

which can be written in the following rigorous way:

$$(A_{\mathcal{K}} v, A_{\mathcal{K}} v) \le c_s \rho_{A_{\mathcal{K}}} ((I - S_{\mathcal{K}})v, v)_{A_{\mathcal{K}}}, \quad \text{for all } v \in \mathcal{K}. \tag{30}$$

Lemma 6.4. *Assume a constrained smoother $R_{\mathcal{K}}$ satisfy (R) and (S_ρ). Then the smoothing property (S_α) holds for $\alpha = 1$.*

Proof. As we mentioned earlier, the symmetry of $R_{\mathcal{K}}$ implies $S_{\mathcal{K}}$ is symmetric in $(\cdot,\cdot)_{A_{\mathcal{K}}}$. The assumption (R) implies $S_{\mathcal{K}}$ is a contraction.

We use formulation (30) of the assumption (S_ρ) and the symmetry of $S_{\mathcal{K}}$ to get

$$(A_{\mathcal{K}} S_{\mathcal{K}}^m v, A_{\mathcal{K}} S_{\mathcal{K}}^m v) \le c_s \rho_{A_{\mathcal{K}}} ((I - S_{\mathcal{K}}) S_{\mathcal{K}}^m v, S_{\mathcal{K}}^m v)_A = c_s \rho_{A_{\mathcal{K}}} ((I - S_{\mathcal{K}}) S_{\mathcal{K}}^{2m} v, v)_A.$$

From the elementary inequality

$$x^{2m} \le \frac{1}{2m} \frac{1 - x^{2m}}{1 - x}, \quad \text{for } x \in [0, 1),$$

we obtain the corresponding operator form

$$((I - S_{\mathcal{K}}) S_{\mathcal{K}}^m u, S_{\mathcal{K}}^m u)_A \le \frac{1}{2m} ((I - S_{\mathcal{K}}^{2m}) u, u)_A.$$

Since $S_{\mathcal{K}}$ is a contraction, $\rho(I - S_{\mathcal{K}}^{2m}) \le 1$, and consequently we get the desirable inequality

$$\|A_{\mathcal{K}} S_{\mathcal{K}}^m v\|^2 \le \frac{c_s}{2m} \rho_{A_{\mathcal{K}}} \|v\|_{A_{\mathcal{K}}}^2. \quad \square \tag{31}$$

As noted in [1], let $v^m = S_{\mathcal{K}}^m v$ and $q = -(BB^T)^{-1} B A v^m$. We can write the norm

$$\|A_{\mathcal{K}} S_{\mathcal{K}}^m v\| = \|Q_{\mathcal{K}} A v^m\| = \|A v^m - B^T (BB^T)^{-1} B A v^m\| = \|A v^m + B^T q\|.$$

Namely we can find a pressure term $q \in \mathcal{P}$ to write the smoothing property in the form

$$\|A v^m + B^T q\|^2 \le \frac{c_s}{2m} \rho_A \|v\|_A^2. \tag{32}$$

The extra pressure term comes naturally in the evaluation of $A_{\mathcal{K}}$.

Theorem 6.5. *Assume a constrained smoother $R_{\mathcal{K}}$ satisfy (R) and (S_ρ). Then the smoothing property (S_α) holds for any $\alpha \in [0, 1]$.*

Proof. For $\alpha = 0$, since $R_{\mathcal{K}}$ satisfy (R), $S_{\mathcal{K}}$ is a contraction and thus (28) holds for $\alpha = 0$. The case $\alpha = 1$ has been proved in Lemma 6.4. Interpolation will lead to the conclusion. $\square$

We shall verify the B–S smoother, which is the Richardson iteration for (5), will satisfy the assumptions (R) and (S_ρ) and consequently prove that B–S smoother satisfies the smoothing property (S_α).

Recall that B–S smoother corresponds to $R_{\mathcal{K}} = \omega^{-1} Q_{\mathcal{K}}$. To satisfy (R), we require $\omega \ge \rho_{A_{\mathcal{K}}}$. A more practical requirement is $\omega \ge \rho_A \ge \rho_{A_{\mathcal{K}}}$ since A, not $A_{\mathcal{K}}$, is explicitly formed. Assumption (S_ρ) will hold with constant $c_s = \omega / \rho_{A_{\mathcal{K}}}$. Therefore we obtain the smoothing property (S_α) for B–S smoother.

Corollary 6.6. *Assume $\rho_{A_{\mathcal{K}}} \le \omega \le c_s \rho_{A_{\mathcal{K}}}$. Then B–S smoother satisfies the smoothing property (S_α) for any $\alpha \in [0, 1]$.*

Remark 6.7. For B–S smoother, $A_{\mathcal{K}} = \omega(I - S_{\mathcal{K}})$ and one can prove a better smoothing property in L^2-norm $\|A_{\mathcal{K}} S_{\mathcal{K}}^m\| = \omega \|(I - S_{\mathcal{K}}) S_{\mathcal{K}}^m\| \le C\omega/m$; see [1]. Here we consider the convergence in A-norm and the best smoothing property (for $\alpha = 1$) is in the order of $1/\sqrt{m}$.

Remark 6.8. For the symmetrized multiplicative Schwarz smoother, it satisfies the assumption (R). But assumption (S_ρ) is unlikely true. Consider smoothers based on a decomposition $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i$ in the finest level with $\dim \mathcal{K}_i = \mathcal{O}(1)$ and $\mathcal{K}_i \subset \mathcal{K}$. Then (S_ρ) would be true if the decomposition is stable in L^2-norm, i.e.

$$\inf_{\sum v_i = v} \sum_{i=1}^N \|v_i\|^2 \lesssim \|v\|^2 \quad \text{for all } v \in \mathcal{K}. \tag{33}$$

For some $v \in \mathcal{K}$, we can find ϕ in an appropriate finite element space such that $v = \operatorname{curl} \phi$. Since $v_i \in \mathcal{K}_i$, it may happen $v_i = \operatorname{curl} \phi_i$, then we would get a stable decomposition of ϕ in H^1-norm. Such H^1 stable decomposition of ϕ does not exist using only subspaces with $\mathcal{O}(1)$ size in one level only.

6.4. Approximation property

We then verify the approximation property using the following partial regularity assumption.

(Reg_α) Let $u \in (H_0^1(\Omega))^d$, $p \in L_0^2(\Omega)$ be the weak solution of Stokes equations (15) with a given data $f \in L^2(\Omega)$. Then $u \in (H^{1+\alpha}(\Omega) \cap H_0^1(\Omega))^d$, $p \in H^\alpha(\Omega)$ and

$$\|u\|_{1+\alpha} + \|p\|_\alpha \lesssim \|f\|_{\alpha-1} \tag{34}$$

holds for some $\alpha \in (1/2, 1]$.

Theorem 6.9. *Assume* (Reg_α) *holds for Stokes equations. Define* $P_H : \mathcal{K} \to \mathcal{K}_H \subset \mathcal{V}$ *as* $P_H = I_H I_{\mathcal{K}_H} A_{\mathcal{K}_H}^{-1} I_{\mathcal{K}_H}^T I_H^T Q_{\mathcal{K}}^T A_{\mathcal{K}}$. *Then the following approximation property holds:*

$$|||(I - P_H)u_h|||_{1-\alpha} \lesssim h^\alpha \|u_h\|_A, \quad \text{for all } u_h \in \mathcal{K}_h. \tag{35}$$

Proof. We denote by (u_h, p_h) the solution to

$$(u_h, v_h)_A + (p_h, Bv_h) = (A_{\mathcal{K}} u_h, v_h) \qquad \text{for all } v_h \in \mathcal{V}_h, \tag{36}$$

$$(Bu_h, q_h) = 0 \qquad \text{for all } q_h \in \mathcal{P}_h. \tag{37}$$

By the inf–sup condition, we can chose $v_h \in \mathcal{V}_h$ such that $Bv_h = p_h$ and $\|v_h\|_A \lesssim \|p_h\|$. Choosing such v_h in (36), we can bound the L^2 norm of p_h as

$$\|p_h\|^2 \le |(Au_h, Q_{\mathcal{K}} v_h)| + |(u_h, v_h)_A| \le \|u_h\|_A \|v_h\|_A \lesssim \|u_h\|_A \|p_h\|,$$

which implies

$$\|p_h\| \lesssim \|u_h\|_A. \tag{38}$$

By definition of P_H, there exists a $p_H \in \mathcal{P}_H$ such that $(P_H u_h, p_H)$ satisfy equations

$$(P_H u_h, v_H)_A + (p_H, Bv_H) = (A_{\mathcal{K}} u_h, v_H) \qquad \text{for all } v_H \in \mathcal{V}_H,$$

$$(BP_H u_h, q_H) = 0 \qquad \text{for all } q_H \in \mathcal{P}_H.$$

Similar to (38), we have $\|p_H\| \lesssim \|P_H u_h\|_A \le \|u_h\|_A$.

We thus obtain the orthogonality

$$((I - P_H)u_h, v_H)_A = -(p_h - p_H, Bv_H), \qquad \text{for all } v_H \in \mathcal{V}_H,$$

$$(B(I - P_H)u_h, q_H) = 0, \qquad \text{for all } q_H \in \mathcal{P}_H.$$

We now estimate the norm $|||(I - P_H)u_h|||_{1-\alpha}$ by the standard duality argument. Let $\rho \in H^{\alpha-1}(\Omega)$ and $\eta \in (H^{1+\alpha}(\Omega) \cap H_0^1(\Omega))^d$, $\chi \in H^\alpha(\Omega)$ satisfy

$$(\eta, v)_A + (\chi, Bv) = (\rho, v) \quad \text{for all } v \in (H_0^1(\Omega))^d, \tag{39}$$

$$(B\eta, q) = 0 \quad \text{for all } q \in L_0^2(\Omega). \tag{40}$$

By the partial regularity assumption (Reg_α), we have

$$\|\eta\|_{1+\alpha} + \|\chi\|_\alpha \lesssim \|\rho\|_{\alpha-1}.$$

We chose $v = (I - P_H)u_h$ in (39), to get, for any $\eta_H \in \mathcal{K}_H$ and $\chi_H \in \mathcal{P}_H$,

$$\begin{aligned}(\rho, (I - P_H)u_h) &= (\eta, (I - P_H)u_h)_A + (\chi, B(I - P_H)u_h)\\ &= (\eta - \eta_H, (I - P_H)u_h)_A + (B(\eta - \eta_H), p_h - p_H) + (\chi - \chi_H, B(I - P_H)u_h)\\ &\lesssim H^\alpha (\|\eta\|_{1+\alpha} + \|\chi\|_\alpha) [\|(I - P_H)u_h\|_A + \|p_h\| + \|p_H\|]\\ &\lesssim h^\alpha \|\rho\|_{\alpha-1} \|u_h\|_A,\end{aligned}$$

which implies

$$\|(I - P_H)u_h\|_{1-\alpha} = \sup_{\rho \in H^{1-\alpha}} \frac{(\rho, (I - P_H)u_h)}{\|\rho\|_{\alpha-1}} \lesssim h^\alpha \|(I - P_H)u_h\|_A \le h^\alpha \|u_h\|_A. \quad \square$$

With the smoothing property (S_α) and the approximation property (A_α) using only partial regularity assumption, we can conclude that the W-cycle multigrid using the B–S smoother is uniform convergent in A-norm if the smoothing steps are sufficiently large. Note that a convergence result of multigrid methods using B–S smoother has been obtained in [1] but in l^2-norm and with the full regularity assumption.

6.5. Variable V-cycle multigrid preconditioner

In the convergence analysis of the two-grid and W-cycle multigrid methods, we require the smoothing steps that are sufficiently large. Now we discuss possible ways of relaxing this requirement.

We can view the operator $Q_{\mathcal{K}} : \mathcal{K}_H \to \mathcal{K}$ as the prolongation operator between the non-nested spaces $\mathcal{K}_H$ and $\mathcal{K}$. Combining the stability $\|Q_{\mathcal{K}}\|_A \leq C$, the smoothing property (S_ρ), the assumption (R), and the approximation property, we can apply the framework developed in [18] for non-nested multigrid methods to conclude that the condition number $\kappa(V_{MG}A_{\mathcal{K}}) \lesssim 1$ when variable V-cycle multigrid V_{MG} is used. Here variable V-cycle refers to a V-cycle with variable number of smoothing steps in each level. Moving from fine to coarse grids, the sequence of smoothing steps is geometrically increasing with a certain factor. A typical choice is: 1, 2, 3, 5, 8, 12, 18 . . . . Variable V-cycle has the same computation complexity as the W-cycle but the error operator of the variable V-cycle can be proven to be SPD while that of W-cycle could be non-SPD [18].

Thus preconditioned conjugate gradient (PCG) method can be used for solving the SPD problem (5); see the formulation in Braess and Dahmen [14]. Note that in the evaluation of $A_{\mathcal{K}}$, an extra L^2-projection is needed.

Acknowledgment

The author would like to thank Dr. Yongke Wu for the careful proof reading and referees for valuable comments. The author was supported by NSF Grant DMS-1418934.

References

[1] D. Braess, R. Sarazin, An efficient smoother for the Stokes equation, Appl. Numer. Math. 23 (1) (1997) 3–19.
[2] J. Schöberl, Multigrid methods for a parameter dependent problem in primal variables, Numer. Math. 84 (1999) 1–19.
[3] S.P. Vanka, Block-implicit multigrid solution of Navier–Stokes equations in primitive variables, J. Comput. Phys. 65 (1) (1986) 138–158.
[4] L. Chen, Multigrid methods for constrained minimization problems and application to saddle point problems (submitted for publication) 2014, pp. 1–22.
[5] J. Schöberl, W. Zulehner, On Schwarz-type smoothers for saddle point problems, Numer. Math. 95 (2003) 377–399.
[6] W. Hackbusch, Multigrid Methods and Applications, in: Computational Mathematics, vol. 4, Springer-Verlag, Berlin, 1985.
[7] R.E. Bank, T. Dupont, An optimal order process for solving elliptic finite element equations, Math. Comp. 36 (1981) 35–51.
[8] J.H. Bramble, J.E. Pasciak, J. Wang, J. Xu, Convergence estimates for multigrid algorithms without regularity assumptions, Math. Comp. 57 (1991) 23–45.
[9] J.H. Bramble, J.E. Pasciak, New estimates for multigrid algorithms including the V-cycle, Math. Comp. 60 (1993) 447–471.
[10] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Rev. 34 (1992) 581–613.
[11] J. Xu, L. Zikatanov, The method of alternating projections and the method of subspace corrections in Hilbert space, J. Amer. Math. Soc. 15 (2002) 573–597.
[12] S. Brenner, Convergence of the multigrid V-cycle algorithm for second-order boundary value problems without full elliptic regularity, Math. Comp. 71 (238) (2002) 507–525.
[13] R. Verfürth, A multilevel algorithm for mixed problems, SIAM J. Numer. Anal. 21 (2) (1984) 264–271.
[14] D. Braess, W. Dahmen, A cascadic multigrid algorithm for the Stokes equations, Numer. Math. 82 (1999) 179–191.
[15] M.A. Olshanskii, Multigrid analysis for the time dependent Stokes problem, Math. Comp. 5718 (2011) 1–23.
[16] W. Zulehner, A class of smoothers for saddle point problems, Computing 65 (3) (2000) 227–246.
[17] S.C. Brenner, H. Li, L.-Y. Sung, Multigrid methods for saddle point problems: Stokes and Lamé systems, Numer. Math. (2014).
[18] J.H. Bramble, J.E. Pasciak, J. Xu, The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms, Math. Comp. 56 (1991) 1–34.
[19] M. Benzi, G.H. Golub, J. Liesen, Numerical solution of saddle point problems, Acta Numer. 14 (2005) 1–137.
[20] K. Mardal, R. Winther, Preconditioning discretizations of systems of partial differential equations, Numer. Linear Algebra Appl. 18 (2011) 1–40.
[21] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer-Verlag, 1991.
[22] M. Griebel, P. Oswald, On the abstract theory of additive and multiplicative Schwarz methods, Numer. Math. 70 (1995) 163–180.
[23] L. Chen, Deriving the X-Z identity from auxiliary space method, in: Y. Huang, R. Kornhuber, O. Widlund, J. Xu (Eds.), Domain Decomposition Methods in Science and Engineering XIX, Springer, Berlin, Heidelberg, 2010, pp. 309–316.
[24] V. Sarin, A. Sameh, An efficient iterative method for the generalized Stokes problem, SIAM J. Sci. Comput. 19 (1) (1998) 206–226.
[25] P.M. Crouzeix, P.A. Raviart, Conforming and nonconforming finite element methods for solving the stationary Stokes equations I, R.A.I.R.O 76 (1973) 3–33.
[26] D.F. Griffiths, Finite elements for incompressible flow, Math. Methods Appl. Sci. 1 (1979) 16–31.
[27] D.F. Griffiths, The construction of approximately divergence-free finite elements, in: Mathematics of Finite Elements and Applications, III Proc. Third MAFELAP Conf., Brunel Univ., Uxbridge, 1978, 1979, pp. 237–245.
[28] X. Ye, C. Hall, A discrete divergence-free basis for finite element methods, Numer. Algorithms 16 (1997) 365–380.
[29] A. Tavakoli, Convergence of multigrid algorithms in divergence free space for Stokes problem, Int. J. Appl. Math. Mech. 6 (3) (2010) 62–76.
[30] P. Oswald, Remarks on multilevel bases for divergence-free finite elements, Numer. Algorithms 27 (2001) 131–152.