15

Rapid #: -1596466 **Ariel IP: 128.112.205.74**

Status	Rapid Code	Branch Name	Start Date
Pending	UBY	Main Library	2/12/2008 10:46:26 AM

CALL #: http://journals.wspc.com.sg/m3as/mkt/archive.shtml

LOCATION: **UBY :: Main Library :: PER**

TYPE: Article CC:CCG

JOURNAL TITLE: Mathematical models and methods in applied sciences

USER JOURNAL TITLE: Mathematical Models and Methods in Applied Sciences (M3AS)

UBY CATALOG TITLE: Mathematical models & methods in applied sciences :

ARTICLE TITLE: Robust subspace correction methods for nearly singular systems

ARTICLE AUTHOR: Lee, Y.-J. , Wu, J. , Xu, J., Zikatanov, L.

VOLUME: 17

ISSUE: 11

MONTH:

YEAR: 2007

PAGES: 1937 - 1963

ISSN: 0218-2025

OCLC #: 49229234

CROSS REFERENCE ID: 381431

VERIFIED:

BORROWER: **PUL :: Interlibrary Services, Firestone**

PATRON: **Chen,Jin**

PATRON ID: -

PATRON ADDRESS: -

PATRON PHONE: -

PATRON FAX: -

PATRON E-MAIL: -

PATRON DEPT: -

PATRON STATUS: -

PATRON NOTES: -

Mathematical Models and Methods in Applied Sciences
Vol. 17, No. 11 (2007) 1937–1963

ROBUST SUBSPACE CORRECTION METHODS FOR NEARLY SINGULAR SYSTEMS

YOUNG-JU LEE

Department of Mathematics
Rutgers, The State University of New Jersey
Hill Center, Piscataway, NJ 08854-8019, USA

JINBIAO WU

Laboratory of Mathematics and Applied Mathematics
School of Mathematical Sciences
Peking University, Beijing 100871, China
jwu@math.pku.edu.cn

JINCHAO XU

Department of Mathematics
The Pennsylvania State University
University Park, PA 16802, USA
and
Laboratory of Mathematics and Applied Mathematics
School of Mathematical Sciences, Peking University
Beijing 100871, China
xu@math.psu.edu

LUDMIL ZIKATANOV

Department of Mathematics
The Pennsylvania State University
University Park, PA 16802, USA
ludmil@psu.edu

Received 8 January 2007
Revised 8 May 2007
Communicated by F. Brezzi

In this paper, we discuss convergence results for general (successive) subspace correction methods for solving *nearly* singular systems of equations. We provide parameter independent estimates under appropriate assumptions on the subspace solvers and space decompositions. The main assumption is that any component in the kernel of the singular part of the system can be decomposed into a sum of local (in each subspace) kernel components. This assumption also covers the case of "hidden" nearly singular behavior due to decreasing mesh size in the systems resulting from finite element discretizations of second order elliptic problems. To illustrate our abstract convergence framework, we analyze a multilevel method for the Neumann problem (H(grad) system), and also two-level methods for H(div) and H(curl) systems.

Keywords: Nearly singular problems; subspace corrections; nonexpansive operators; multigrid, domain decomposition.

AMS Subject Classification: 47A58, 47N10, 65N22, 65N55

1. Introduction

We consider the solution of the nearly singular system of equations: Find $u \in V$ such that

$$Au = (A_0 + \epsilon A_1)u = f, \tag{1.1}$$

where V is a finite dimensional Hilbert space, A_0 is symmetric and positive semi-definite, and A_1 is symmetric and positive definite, and $\epsilon > 0$ is a parameter. We are interested in the convergence properties of a special class of linear iterative methods (subspace correction methods) for the approximate solution of (1.1). For $\epsilon = 0$, the operator A reduces to A_0, which is symmetric and positive semi-definite and a convergence study in this case has been done in Ref. 26. Our considerations here are for $\epsilon > 0$, and we focus on proving ϵ-independent convergence results under appropriate (minimal) assumptions.

There are abundant examples that fall into the category of nearly singular systems like (1.1). Such examples are given by the finite element discretizations for H(grad), H(div), and H(curl) systems as discussed in Refs. 23, 3 and 1; stable discretizations of the nearly incompressible linear elasticity problems (see Refs. 33, 31 and 36),

$$-\nabla(\operatorname{div} u) - (1 - 2\nu)\Delta u = f \quad \text{in } \Omega, \tag{1.2}$$

where the Poisson's ratio ν is close to $1/2$. Nearly singular problems also occur when solving indefinite systems arising from mixed finite element discretizations such as (Navier–) Stokes equations[20] or more complicated system of equations such as non-Newtonian flow equations as discussed in Refs. 6, 24 and 27 by augmented Lagrangian method (see Refs. 21, 13 and 19).

The main goal in this paper is to provide unified theoretical framework for the analysis of multilevel methods for nearly singular systems of equations. Apparently, problems like (1.1) are badly conditioned, since the smaller ϵ is the larger the condition number of A becomes. To illustrate how this affects the convergence rate of an iterative method, let us consider Gauss–Seidel iterations, applied to the following simple example:

$$Au = (A_0 + \epsilon A_1)u = \left\{ \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} + \epsilon \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\} u = f. \tag{1.3}$$

Table 1. The number of iterations to obtain the energy norm error $\|u - u^\ell\|_A < 10^{-6}$ for various values of ϵ.

ϵ	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	0
# of iterations	14	94	823	7427	66,556	588,770	2

Here, we assume that $f = (f_1, f_2, f_3)^t \in \mathbb{R}^3$ and f is in the range of A_0. The convergence history of the Gauss–Seidel method for (1.3) for different values of ϵ is recorded in Table 1. It is clearly seen that for ϵ vanishingly small, the convergence deteriorates. To explain such behavior, we consider the energy norm convergence rate δ of the Gauss–Seidel method for (1.3), which is given by (see Refs. 35 and 26),

$$\delta^2 = 1 - \frac{1}{K}, \quad K = 1 + \sup_{v=(v_1,v_2,v_3)^t \in \mathbb{R}^3} \frac{(1+\epsilon)^{-1} v_2^2 + (2+\epsilon)^{-1} v_3^2}{(Av, v)}. \tag{1.4}$$

Clearly, by choosing v in the null space of A_0, for example, $v = (1,1,1)^t$, we obtain

$$K \geq 1 + \frac{(1+\epsilon)^{-1} + (2+\epsilon)^{-1}}{3\epsilon}. \tag{1.5}$$

Hence $K \to \infty$ and $\delta \to 1$, when $\epsilon \to 0$.

Of course, the Gauss–Seidel method is a special iterative method, and if we consider the Conjugate Gradient or the Richardson method, the convergence behavior will be different. Indeed, in the extreme case when the right-hand side f and the initial guess u^0 belong to the range of A_0 all the iterates u^ℓ generated by Richardson method or the Conjugate Gradient method belong to the range of A_0. This is due to the fact that in the above example, $A_1 = I$, i.e. $A = A_0 + \epsilon I$. In such extreme case, with special initial guess, it is not difficult to show that the convergence rate of Richardson method or the Conjugate Gradient method is independent of the parameter ϵ. However, most subspace correction methods including multigrid methods do not in general possess this property. Moreover, for a practical iterative method, it is desirable that ϵ-uniform bound on the convergence holds for *any initial guess* u^0, which rules out the special choice $u^0 \in \mathcal{R}(A_0)$. Generally speaking, if we decompose the solution u to (1.1) as $u = v + c$ where v and c belong to the range and the null space of A_0, respectively, the difficulty is to approximate c (the component of u in the null space of A_0). In summary, from the results in Table 1 and (1.5), we may conclude that a naive application of an algorithm based on space decomposition and subspace corrections (such as Gauss–Seidel relaxation) would result in an inefficient method.

As it turns out, a crucial assumption on the space decomposition is needed, in order to obtain ϵ-uniform convergence (see Sec. 4.2, assumption (**A1**)). Roughly speaking, this assumption says that every kernel component of A_0 can be decomposed as a sum of "local" kernel components in the subspaces. For example, for finite element discretizations of variational problems in $H(\text{curl})$ and $H(\text{div})$, pointwise smoothers violates (**A1**), while block Schwarz smoothers with blocks corresponding to vertices (or edges in $H(\text{div})$) satisfy (**A1**). In the former case (pointwise smoothers), it can be shown that a two-level method is not optimal (see Ref. 37). In the latter case, when using block smoothers, a uniform convergence is proved in Refs. 23 and 3. These results show that (**A1**) is necessary and sufficient for ϵ-uniform convergence.

The rest of the paper is organized as follows. In Sec. 2 we introduce some of the frequently used notation. In Sec. 2.1, we discuss examples of nearly singular problems, and in Sec. 2.2 we show how to reduce an indefinite problem to a nearly singular problem. In Sec. 3 we first formulate a method based on the augmented system and analyze the simplest case when the kernel of A_0 is one dimensional. In Sec. 4 we present the subspace correction algorithm, the main assumptions, and the abstract convergence theory. Further, in Sec. 4.1 we also prove that the augmented system methods are equivalent to subspace correction methods and hence our analysis applies in more general cases, when the null space of A_0 is not just one dimensional. To illustrate the abstract results, in Sec. 5 we derive convergence estimates for a multilevel method for the Neumann problem and for two-grid methods for variational problems in $H(\text{div})$ and $H(\text{curl})$. Finally in Sec. 6 we give some concluding remarks.

2. Notation and Preliminaries

Throughout this paper, we use the notation introduced below. Let V be a finite dimensional Hilbert space with an inner product $(\cdot,\cdot)$ and a corresponding norm $\|\cdot\|$. For a bounded operator $T : V \mapsto V$, $\mathcal{N}(T)$ and $\mathcal{R}(T)$ denote the null space of T and the range of T respectively. Let $\Omega \subset \mathbb{R}^d, d = 2, 3$ be a connected bounded domain. We will use the standard notation for the differential operators grad, div, and curl. The space $L^2(\Omega)$ denotes the space of square integrable functions and $H^1(\Omega) = H(\text{grad})$ denotes the standard Sobolev space consisting of square integrable functions with square integrable (weak) derivatives of first order. Similarly, $H(\text{curl})$ and $H(\text{div})$ denote the spaces of $L^2(\Omega)$ functions with square integrable curl or divergence, respectively. The bilinear forms $(\cdot,\cdot)_0$ and $(\cdot,\cdot)_1$ denote the usual $L^2(\Omega)$ inner product and $H^1(\Omega)$ inner product, respectively. Also $\|\cdot\|_0$, $\|\cdot\|_1$, and $|\cdot|_1$ denote the L^2 norm, H^1 norm, and H^1 semi-norm, respectively. We also denote $\|\cdot\|_{\text{div}}$, $|\cdot|_{\text{div}}$, $\|\cdot\|_{\text{curl}}$, and $|\cdot|_{\text{curl}}$ by $H(\text{div})$ norm, $H(\text{div})$ semi-norm, $H(\text{curl})$ norm, and $H(\text{curl})$ semi-norm, respectively. Following Ref. 34, we use the notation $x_1 \lesssim y_1$ and $x_2 \gtrsim y_2$ whenever there exist constants C_1 and C_2 independent of ϵ and other important parameters such that $x_1 \leq C_1 y_1$ and $x_2 \geq C_2 y_2$.

2.1. *Examples of positive definite nearly singular problems*

We summarize here some examples that fall into the category of nearly singular problems and in our opinion, are encountered quite often in the numerical models of physical phenomena.

2.1.1. *Discretizations of variational problems in* $H(\mathrm{grad})$, $H(\mathrm{curl})$ *and* $H(\mathrm{div})$

For this type of problems we have

$$(A_0 u, v) := (Gu, Gv), \quad (\epsilon A_1 u, v) := (\tau u, v),$$

where $G = \mathrm{curl}$ or $G = \mathrm{div}$ and τ is a given function. As we mentioned in Sec. 1, the nearly singular behavior is "hidden" in the lower order term. In Secs. 5.1 and 5.2, we apply the abstract theory to analyze the convergence of two-level methods for variational problems in $H(\mathrm{curl})$ and $H(\mathrm{div})$, and multilevel method for variational problem in $H(\mathrm{grad})$.

2.1.2. *Anisotropic problems*

An interesting example, which we will not consider further in this paper is related to the finite element discretizations of anisotropic problems:

$$(Au, v) := (a\,\mathrm{grad}\,u,\ \mathrm{grad}\,v),$$

and $V = H_0^1(\Omega)$. The coefficient tensor a behaves differently in different coordinate directions, and we may take $a = \mathrm{diag}(1, 1, \ldots, \epsilon)$, for $0 < \epsilon \le \epsilon_{\max} < 1$. Apparently (Au, v) can then be rewritten as $A = A_0 + \epsilon A_1$, where

$$(A_0 u, v) := \sum_{i=1}^{d-1}((1-\epsilon)\partial_i u, \partial_i v), \quad (\epsilon A_1 u, v) := (\epsilon\,\mathrm{grad}\,u,\ \mathrm{grad}\,v).$$

We remark that when the mesh is aligned with the anisotropy and the null space of A_0 is known, there are several techniques that can be applied to design of an optimal iterative method. These techniques use semi-coarsening, line (plane) smoothers or both and they precisely correspond to splittings that satisfy assumption (**A1**) (see Sec. 4.2). When the mesh is not aligned with the anisotropy or a varies throughout the computational domain, to find a splitting that satisfies (**A1**) is a challenging and complicated task, which we will study in the future.

2.2. *Reduction of indefinite problems to nearly singular problems*

Other examples of nearly singular problems are obtained when solving systems arising from mixed and hybrid finite element discretizations of second order partial differential equations,[13] such as the indefinite systems corresponding to mixed finite element discretizations for elliptic problems (Darcy's law in porous media) or the Stokes equation.[7,14,30] We aim to show that implementing an efficient iterative method for the resulting indefinite linear system reduces to designing an efficient method for the solution of an auxiliary nearly singular problem.

To begin, we briefly describe a minimal set-up needed for our discussion. Let V and W be two Hilbert spaces, and V^* and W^* be the dual spaces of V and

W, respectively. Here $\langle \cdot, \cdot \rangle$ will denote the dual pairing. Consider the following variational problem: Find $(u, p) \in V \times W$ such that

$$\begin{cases} Au + B^* p = f, \\ Bu = g, \end{cases} \tag{2.1}$$

where $f \in V^*$, $g \in W^*$, $A : V \mapsto V^*$, $B : V \mapsto W^*$, and $B^* : W \mapsto V^*$ is the dual operator of B, namely

$$\langle v, B^* q \rangle = \langle q, Bv \rangle \quad \forall\ v \in V, \ \ q \in W. \tag{2.2}$$

We assume that A is coercive, B is onto and B satisfies the following inf-sup condition:

$$\langle Av, v \rangle \ \ \geq \ \ \alpha \|v\|_V^2, \quad \alpha > 0, \quad \inf_{q \in W} \sup_{v \in V} \frac{\langle q, Bv \rangle}{\|v\|_V \|q\|_W} = \beta > 0.$$

By the Babuška–Brezzi theory,[3,12,13] problem (2.1) is well-posed. Since same conditions would hold for any compatible discretization of (2.1), we may assume here, without loss of generality, that V and W is a pair of compatible finite element spaces. One of the most popular methods for solving discrete equations resulting from (2.1) is the classical Uzawa method: For given (u^ℓ, p^ℓ), the next iterate $(u^{\ell+1}, p^{\ell+1})$ is obtained by

$$Au^{\ell+1} = f - B^* p^\ell, \qquad p^{\ell+1} = p^\ell - \kappa(g - Bu^{\ell+1}),$$

where $\kappa > 0$ is a damping parameter. It is known that the Uzawa method converges if κ satisfies $0 < \kappa < \frac{2}{\rho(S)}$, where $S = BA^{-1}B^*$ is the Schur complement corresponding to the p variable. One may say that the Uzawa method is equivalent to a Richardson relaxation for a linear system with the Schur complement S. Certainly, the choice of the damping parameter κ affects the convergence of the Uzawa's iteration and to come up with a procedure for choosing optimal damping parameter is a non-trivial task (if at all possible). One remedy is to use the Augmented Lagrangian method,[13,19,21] which solves a modified problem, equivalent to (2.1) by the Uzawa method. The modified problem is as follows:

$$\begin{cases} (A + \epsilon^{-1} B^* B)u + B^* p = f + \epsilon^{-1} B^* g, \\ Bu = g. \end{cases} \tag{2.3}$$

Application of the Uzawa method to (2.3) with damping parameter $\kappa = \epsilon^{-1}$ reads: Given (u^ℓ, p^ℓ), the new iterate $(u^{\ell+1}, p^{\ell+1})$ is obtained by solving the following system:

$$\begin{cases} (A + \epsilon^{-1} B^* B)u^{\ell+1} = f + \epsilon^{-1} B^* g - B^* p^\ell, \\ p^{\ell+1} = p^\ell - \epsilon^{-1}(g - Bu^{\ell+1}). \end{cases} \tag{2.4}$$

Convergence of this method has been discussed in several works.[9,19,21] Here, we also provide a result on the rate of convergence of the augmented Lagrangian method, which indicates that for small ϵ, the iterates converge very fast to the solution of (2.1). The payoff, however, is that to obtain $u^{\ell+1}$ we need to solve a nearly singular system, such as

$$(\epsilon A + B^*B)u^{\ell+1} = \epsilon f + B^*g - \epsilon B^*p^\ell. \tag{2.5}$$

In many applications, B^*B would be a *singular* operator. Thus, the design of optimal iterative method for the nearly singular system (2.5) will result in an optimal iterative method for the indefinite system (2.1) and a general solution strategy for various mixed and hybrid finite element discretizations.

We now state and prove a convergence result for the augmented Lagrangian method.

Lemma 2.1. *Let (u^0, p^0) be a given initial guess and for $\ell \geq 1$, let (u^ℓ, p^ℓ) be the iterates obtained via the augmented Lagrangian method. Then the following estimates hold:*

$$\|p - p^\ell\|_W \leq \left(\frac{\epsilon}{\epsilon + \mu_0}\right)^\ell \|p - p^0\|_W,$$

$$\|u - u^\ell\|_A \leq \sqrt{\epsilon}\|p - p^{\ell-1}\|_W \leq \sqrt{\epsilon}\left(\frac{\epsilon}{\epsilon + \mu_0}\right)^\ell \|p - p^0\|_W,$$

where μ_0 is the minimum eigenvalue of $S = BA^{-1}B^*$.

Proof. Let $e_p^\ell = (p - p^\ell)$ and $e_u^\ell = (u - u^\ell)$ be the errors after ℓ iterations ($\ell > 0$). From the definition of (u^ℓ, p^ℓ), (2.3), and (2.4) we obtain that (e_u^ℓ, e_p^ℓ) satisfy:

$$\begin{cases} (A + \epsilon^{-1}B^*B)e_u^\ell = -B^*e_p^{\ell-1}, \\ e_p^\ell = (I - \epsilon^{-1}B(A + \epsilon^{-1}B^*B)^{-1}B^*)e_p^{\ell-1}. \end{cases}$$

A simple application of Shermann–Morrison–Woodburry formula for $(\epsilon A + B^*B)^{-1}$ gives

$$\epsilon^{-1}B(A + \epsilon^{-1}B^*B)^{-1}B^* = B(\epsilon A + B^*B)^{-1}B^* = S_\epsilon - S_\epsilon(I + S_\epsilon)^{-1}S_\epsilon,$$

where $S_\epsilon = \epsilon^{-1}S = \epsilon^{-1}BA^{-1}B^*$. It is also straightforward to verify that

$$I - \epsilon^{-1}B(A + \epsilon^{-1}B^*B)^{-1}B^* = I - S_\epsilon + S_\epsilon(I + S_\epsilon)^{-1}S_\epsilon = (I + S_\epsilon)^{-1}. \tag{2.6}$$

Hence

$$e_p^\ell = (I + \epsilon^{-1}S)^{-1}e_p^{\ell-1}$$

and the first estimate follows immediately. The second estimate is obtained using the following identities:

$$\begin{aligned} \|e_u^\ell\|_A^2 &= (Ae_u^\ell, e_u^\ell) = ((A + \epsilon^{-1}B^*B - \epsilon^{-1}B^*B)e_u^\ell, e_u^\ell) \\ &= -(B^*e_p^{\ell-1}, e_u^\ell) - (\epsilon^{-1}Be_u^\ell, Be_u^\ell) \\ &= \epsilon[(Ze_p^{\ell-1}, e_p^{\ell-1}) - \|Ze_p^{\ell-1}\|_W^2], \end{aligned}$$

where $Z = B(\epsilon A + B^*B)^{-1}B^*$. The proof is complete by observing that (2.6) implies that $\|Z\|_W \leq 1$. □

3. Solution of Nearly Singular Problems via Augmented System

The first method for nearly singular problems that we now present in this paper is based on the so-called augmented or modified system of equations[4,29] by means of a basis of $\mathcal{N}(A_0)$ is known. We equip $V = \mathbb{R}^n$ with the discrete ℓ^2 inner product $(\cdot,\cdot) = (\cdot,\cdot)_{\ell^2}$ and a norm $\|\cdot\| = (\cdot,\cdot)^{1/2}$. Assuming that a basis in the kernel of A_0 is known:

$$\mathcal{N}(A_0) \subset \operatorname{span}\{\phi_1, \ldots, \phi_m\}. \tag{3.1}$$

We define $\Phi = [\phi_1, \ldots, \phi_m] : W \mapsto V$ with $W = \mathbb{R}^m$ and obtain the augmented system of equations for the problem (1.1) as follows (see Refs. 4 and 29):

$$\mathcal{A}\widetilde{u} = \begin{pmatrix} \Phi^t A \Phi & \Phi^t A \\ A\Phi & A \end{pmatrix} \widetilde{u} = \widetilde{f}, \qquad \widetilde{f} = \begin{pmatrix} \Phi^t f \\ f \end{pmatrix}. \tag{3.2}$$

We note that the system (3.2) is singular and, furthermore, the range and null space of $\mathcal{A}$ are characterized by

$$\mathcal{R}(\mathcal{A}) = \left\{ \begin{pmatrix} \Phi^t v \\ v \end{pmatrix} : v \in V \right\} \quad \text{and} \quad \mathcal{N}(\mathcal{A}) = \left\{ \begin{pmatrix} c \\ -\Phi c \end{pmatrix} : c \in W \right\}. \tag{3.3}$$

Obviously, there exist infinitely many solutions to Eq. (3.2). However, if $\widetilde{u} = \binom{u_1}{u_2}$ is a solution to (3.2) then the solution u to the original system (1.1) can be recovered in a unique fashion by setting $u = \Phi u_1 + u_2$.

We now consider the block Gauss–Seidel method for the augmented system (3.2) with one block given by $\Phi^t A \Phi$. This method has convergence rate independent of ϵ and it is, as we shall see in Sec. 4.1, equivalent to a two-level method with coarse space $\mathcal{N}(A_0)$.[22,34]

As an example, we first analyze a simpler problem, with $A_1 = I$, and $f \in \mathcal{R}A_0$, that is

$$Au = (A_0 + \epsilon I)u = f, \tag{3.4}$$

where $\dim \mathcal{N}(A_0) = 1$, i.e. $\mathcal{N}(A_0) = \operatorname{span}\{\xi\}$. The augmented system of equations for this problem is given by

$$\mathcal{A}\widetilde{u} = \begin{pmatrix} \xi^t A \xi & \xi^t A \\ A\xi & A \end{pmatrix} \widetilde{u} = \widetilde{f}, \tag{3.5}$$

where $\widetilde{f} = \binom{(\xi,f)}{f} = \binom{0}{f}$.

Assume that

$$A = D_\epsilon - L - L^t, \quad A_0 = D_0 - L - L^t, \tag{3.6}$$

where D_ϵ and D_0 are the diagonal matrices corresponding to A and A_0, respectively, and $-L$ is the strict lower triangle of both A and A_0. We then define

$$S_\epsilon = L D_\epsilon^{-1} L^t \quad \text{and} \quad S_0 = L D_0^{-1} L^t.$$

Since $D_\epsilon = D_0 + \epsilon I$, there exists $\alpha(\epsilon)$ such that for any $v \in V$,

$$(S_0 v, v) \geq (S_\epsilon v, v) \geq \alpha(\epsilon)(S_0 v, v), \quad \text{and} \quad \lim_{\epsilon \to 0} \alpha(\epsilon) = 1. \tag{3.7}$$

We also introduce a splitting of the augmented matrix $\mathcal{A}$

$$\mathcal{A} = \mathcal{D} - \mathcal{L}^t - \mathcal{L} \quad \text{and} \quad \mathcal{S} = \mathcal{L}\mathcal{D}^{-1}\mathcal{L}^t.$$

Note that the null space and the range of $\mathcal{A}$ are

$$\mathcal{N}\mathcal{A} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -\xi \end{pmatrix} \right\} \quad \text{and} \quad \mathcal{R}\mathcal{A} = \left\{ \widetilde{v} = \begin{pmatrix} (\xi, v) \\ v \end{pmatrix} : v \in V \right\}.$$

Furthermore, both the energy norm convergence rate $\delta_{\mathcal{A}}$ of the Gauss–Seidel method for matrix $\mathcal{A}$ and δ_{A_0} of the Gauss–Seidel method for the matrix A_0 are as follows[35,26]:

$$\delta_{\mathcal{A}}^2 = 1 - \frac{1}{K(\mathcal{A})}, \quad K(\mathcal{A}) = 1 + \sup_{\widetilde{v} \in \mathcal{R}\mathcal{A}} \inf_{\widetilde{c} \in \mathcal{N}\mathcal{A}} \frac{(\mathcal{S}(\widetilde{v} + \widetilde{c}), (\widetilde{v} + \widetilde{c}))}{(\widetilde{v}, \widetilde{v})_{\mathcal{A}}}$$

and

$$\delta_{A_0}^2 = 1 - \frac{1}{K(A_0)}, \quad K(A_0) = 1 + \sup_{v \in \mathcal{R}A_0} \inf_{c \in \mathcal{N}A_0} \frac{(S_0(v + c), (v + c))}{(v, v)_{A_0}}.$$

The constant $K(A_0)$ is independent of ϵ and we have the following convergence result.

Lemma 3.1. *Let $\delta_{\mathcal{A}}$ and δ_{A_0} be defined as above. Then*

$$\lim_{\epsilon \to 0} \delta_{\mathcal{A}} = \delta_{A_0}. \tag{3.8}$$

Proof. A simple calculation yields

$$K(\mathcal{A}) = 1 + \sup_{\widetilde{v} = ((\xi, v), v^t)^t \in \mathcal{R}\mathcal{A}} \inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v + \lambda\xi), (v + \lambda\xi)) + \|P_\xi(v + \lambda\xi)\|_A^2}{\|\widetilde{v}\|_{\mathcal{A}}^2}, \tag{3.9}$$

where P_ξ is the A-orthogonal projection on $\mathcal{N}(A_0)$, namely

$$P_\xi = \xi(\xi^t A \xi)^{-1} \xi^t A.$$

Now, for a given $v \in V$, we consider the orthogonal decomposition

$$v = v^0 + \gamma\xi,$$

where $v^0 \in \mathcal{R}A_0$ and $\gamma \in \mathbb{R}$. With this decomposition, $\|\widetilde{v}\|_{\mathcal{A}}^2$ can be written as

$$\|\widetilde{v}\|_{\mathcal{A}}^2 = \|v + (\xi, v)\xi\|_A^2 = \|v^0\|_A^2 + \epsilon|\gamma + (\xi, v)|^2\|\xi\|^2. \tag{3.10}$$

From (3.9), we see that

$$\begin{aligned} K(\mathcal{A}) &\geq 1 + \sup_{v^0 \in \mathcal{R}(A_0)} \sup_{\gamma \in \mathbb{R}} \inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v^0 + \lambda\xi), (v^0 + \lambda\xi))}{\|v^0\|_A^2 + \epsilon|\gamma + (\xi, v)|^2\|\xi\|^2} \\ &\geq 1 + \sup_{v^0 \in \mathcal{R}(A_0)} \inf_{\lambda \in \mathbb{R}} \frac{(S_0(v^0 + \lambda\xi), (v^0 + \lambda\xi))}{\|v^0\|_A^2}. \end{aligned} \tag{3.11}$$

Observe that for $v \in \mathcal{N}(A_0)$, we have

$$\inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v+\lambda\xi),(v+\lambda\xi)) + \|P_\xi(v+\lambda\xi)\|_A^2}{\|\widetilde{v}\|_A^2} = 0$$

and also for $v = v^0 + \gamma\xi \in V$,

$$\begin{aligned}\inf_{\lambda \in \mathbb{R}} &(S_\epsilon(v+\lambda\xi),(v+\lambda\xi)) + \|P_\xi(v+\lambda\xi)\|_A^2 \\ &= \inf_{\lambda \in \mathbb{R}} (S_\epsilon(v^0+\lambda\xi),(v^0+\lambda\xi)) + \|P_\xi(v^0+\lambda\xi)\|_A^2.\end{aligned}$$

Therefore, $K(\mathcal{A})$ can be estimated by

$$\begin{aligned}K(\mathcal{A}) &= 1 + \sup_{\widetilde{v} \in \chi} \inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v+\lambda\xi),(v+\lambda\xi)) + \|P_\xi(v-\lambda\xi)\|_A^2}{\|v^0\|_A^2 + \epsilon|\gamma + (\xi, v)|^2\|\xi\|^2} \\ &\leq 1 + \sup_{\widetilde{v} \in \chi} \inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v+\lambda\xi),(v+\lambda\xi)) + \|P_\xi(v+\lambda\xi)\|_A^2}{\|v^0\|_{A_0}^2}, \\ &= 1 + \sup_{v \in \mathcal{R}(A_0)} \inf_{\lambda \in \mathbb{R}} \frac{(S_\epsilon(v^0+\lambda\xi),(v^0+\lambda\xi)) + \|P_\xi(v^0+\lambda\xi)\|_A^2}{\|v^0\|_{A_0}^2}, \end{aligned} \qquad (3.12)$$

where

$$\chi = \left\{ \begin{pmatrix} (\xi, v) \\ v \end{pmatrix} \in \mathcal{R}(\mathcal{A}) : v = v^0 + \gamma\xi \text{ with } v^0 \neq 0 \in \mathcal{R}(A_0) \right\}.$$

Finally, we observe that

$$\|P_\xi(v^0+\lambda\xi)\|_A^2 = \epsilon\|P_\xi(v^0+\lambda\xi)\|^2. \qquad (3.13)$$

As a result, we see that

$$K(\mathcal{A}) \leq \alpha(\epsilon)K(A_0) + \sup_{v^0 \in \mathcal{R}(A_0)} \inf_{\lambda \in \mathbb{R}} \frac{\epsilon\|P_\xi(v^0+\lambda\xi)\|^2}{\|v^0\|_{A_0}^2}. \qquad (3.14)$$

Thanks to the inequalities (3.11) and (3.14), by taking the limit $\epsilon \to 0$, we complete the proof. □

More general cases can be treated using the analysis provided in Ref. 26, since in Sec. 4.1 we prove that an iterative method for the augmented system is in fact equivalent to a subspace correction method for the original problem (1.1).

4. Abstract Convergence Analysis for General Subspace Correction Method

Let us introduce the bilinear forms that correspond to the operators (matrices) A, A_0, and A_1. We define a, a_0, and $a_1 : V \times V \mapsto \mathbb{R}$ by

$$a(u,v) = (Au, v), \quad a_0(u,v) = (A_0u, v), \quad \text{and} \quad a_1(u,v) = (A_1u, v), \quad \forall\, u, v \in V. \qquad (4.1)$$

The corresponding induced (semi)-norms are then defined by $\|u\|_a^2 = a(u,u)$, $|u|_{a_0}^2 = a_0(u,u)$, and $\|u\|_{a_1}^2 = a_1(u,u)$. The solution $u \in V$ to the system (1.1) satisfies the following variational problem:

$$a(u,v) = a_0(u,v) + \epsilon\, a_1(u,v) = (f,v), \quad \forall\, v \in V. \tag{4.2}$$

The null space of the operator A_0 denoted by $\mathcal{N}$ is given by

$$\mathcal{N} = \{u \in V : a_0(u,v) = 0,\ \forall\, v \in V\}.$$

We then define the $a(\cdot,\cdot)$-orthogonal complement $\mathcal{N}^\perp$ of $\mathcal{N}$:

$$\mathcal{N}^\perp = \{u \in V : a(u,v) = 0,\ \forall v \in \mathcal{N}\} = \{u \in V : a_1(u,v) = 0,\ \forall v \in \mathcal{N}\}.$$

Note that, $\mathcal{N}^\perp$ is also the orthogonal complement of $\mathcal{N}$ with respect to the inner product $a_1(\cdot,\cdot)$. To describe the iterative algorithm that we study, we assume that V is decomposed into a sum of subspaces $\{V_i\}_{i=1}^J$ such that $V = \sum_{i=1}^J V_i$. A subspace correction algorithm then can be written as follows.

Algorithm 4.1 (MSSC). Let $u^0 \in V$ be given,
for $l = 1, \ldots$ until convergence,
$u_0^{l-1} = u^{l-1}$
for $i = 1, \ldots, J$
Let $e_i \in V_i$ solve

$$a^i(e_i, v_i) = f(v_i) - a(u_{i-1}^{l-1}, v_i), \quad \forall v_i \in V_i \tag{4.3}$$

$u_i^{l-1} = u_{i-1}^{l-1} + e_i$
endfor
$u^l = u_J^{l-1}$
endfor

Here, as usual, the bilinear forms $a^i(\cdot,\cdot)$ are approximations of $a(\cdot,\cdot)$ on V_i. For each $i = 1, \ldots, J$, we assume that they satisfy the inf-sup conditions:

$$\inf_{v_i \in V_i} \sup_{w_i \in V_i} \frac{a^i(v_i, w_i)}{\|v_i\| \|w_i\|} = \inf_{w_i \in V_i} \sup_{v_i \in V_i} \frac{a^i(v_i, w_i)}{\|v_i\| \|w_i\|} = \beta_i, \tag{4.4}$$

where $\beta_i > 0$. The assumptions make all the steps in Algorithm 4.1 well-defined.

Remark 4.1. In what follows, we shall often use the following convention: A decomposition $\{V_i\}_{i=1}^J$ such that $\sum_{i=1}^J V_i = V$, and approximating bilinear forms $a^i(\cdot,\cdot)$, satisfying (4.4) uniquely determine all steps in a subspace correction algorithm, such as Algorithm 4.1.

To obtain an estimate on the convergence rate we need to introduce some additional notions and terminology, commonly used in the analysis of iterative methods such as Algorithm 4.1. We first define the subspace solvers $T_i : V \mapsto V_i$, by setting $T_i v$ be the unique solution to the variational problem:

$$a^i(T_i v, v_i) = a(v, v_i), \quad \forall\, v \in V, \quad \forall\, v_i \in V_i. \tag{4.5}$$

We note that if $a^i = a$ on V_i, then T_i is the orthogonal projection on V_i with respect to the inner product $a(\cdot,\cdot)$. In such case we denote the *exact* subspace solver by P_i, instead of T_i. Evidently, after l iterations the errors $(u - u^l)$, $(u - u^{l-1})$, and $(u - u^0)$ are related as follows:

$$\begin{aligned} u - u^l &= E(u - u^{l-1}) \\ &= \cdots = E^l(u - u^0), \quad \text{where } E = (I - T_J)(I - T_{J-1}) \cdots (I - T_1). \end{aligned} \tag{4.6}$$

The operator E is usually referred to as the error transfer operator. For each $1 \le i \le J$, there exists a unique $T_i^* : V \mapsto V_i$, the Hilbert *adjoint* of T_i with respect to the inner product $a(\cdot,\cdot)$, i.e.

$$a(T_i v, w) = a(v, T_i^* w), \quad \forall\, v, w \in V. \tag{4.7}$$

The *symmetrization* $\bar{T}_i$ of T_i for each $i = 1, \ldots, J$ is given by

$$\bar{T}_i = T_i + T_i^* - T_i^* T_i. \tag{4.8}$$

For each $i = 1, \ldots, J$, we introduce a projection $P_{i,1}$. $P_{i,1} : V \mapsto V_i$ defined by

$$a_1(P_{i,1} v, v_i) = a_1(v, v_i), \quad \forall v \in V, \;\; v_i \in V_i.$$

We also need a projection $P_{i,0}$, orthogonal with respect to $a_0(\cdot,\cdot)$. Since this bilinear form has a non-trivial kernel, we need to define the action $P_{i,0}$ in a non-ambiguous way. This can be done as follows. Denote by $\mathcal{N}_i$, the local null space of a_0, namely,

$$\mathcal{N}_i = \{u_i \in V_i : a_0(u_i, v_i) = 0, \; \forall\, v_i \in V_i\}$$

and the orthogonal complement of $\mathcal{N}_i$ with respect to the inner product $a(\cdot,\cdot)$,

$$\mathcal{N}_i^{\perp} = \{u_i \in V_i : a(u_i, v_i) = 0, \; \forall\, v_i \in \mathcal{N}_i\} = \{u_i \in V_i : a_1(u_i, v_i) = 0, \; \forall\, v_i \in \mathcal{N}_i\}.$$

Then $P_{i,0} v \in \mathcal{N}_i^{\perp}$, for all $v \in V$ is uniquely determinated by the following variational problem:

$$a_0(P_{i,0} v, v_i) = a_0(v, v_i), \quad \forall\, v \in V, \quad v_i \in V_i.$$

4.1. *Equivalence with methods for the augmented system*

We now show that every subspace correction method (SSCA) for the augmented system (3.2) corresponds to an equivalent subspace correction method (SSCO) for the original problem (1.1). Let $\widetilde{V} = \mathbb{R}^{n+m}$ and let $\widetilde{a}(\cdot,\cdot) : \widetilde{V} \times \widetilde{V} \mapsto \mathbb{R}$ be the bilinear form corresponding to the augmented matrix $\mathcal{A} : \widetilde{V} \mapsto \widetilde{V}$, namely,

$$\widetilde{a}(\widetilde{u}, \widetilde{v}) = (\mathcal{A}\widetilde{u}, \widetilde{v}), \quad \forall\, \widetilde{u}, \widetilde{v} \in \widetilde{V}. \tag{4.9}$$

In order to introduce the subspace correction method for the augmented system, we decompose the space $\widetilde{V}$ as follows:

$$\widetilde{V} = \sum_{i=1}^{J} \widetilde{V}_i. \tag{4.10}$$

By the convention that we have made above, associated with each subspace $\widetilde{V}_i$, we have a bilinear form $\widetilde{a}^i(\cdot,\cdot) : \widetilde{V}_i \times \widetilde{V}_i \mapsto \mathbb{R}$ satisfying the inf–sup conditions, required for the well-posedness of the local problems:

$$\inf_{\widetilde{v}_i \in \widetilde{V}_i} \sup_{\widetilde{w}_i \in \widetilde{V}_i} \frac{\widetilde{a}^i(\widetilde{v}_i, \widetilde{w}_i)}{\|\widetilde{v}_i\|\|\widetilde{w}_i\|} = \inf_{\widetilde{w}_i \in \widetilde{V}_i} \sup_{\widetilde{v}_i \in \widetilde{V}_i} \frac{\widetilde{a}^i(\widetilde{v}_i, \widetilde{w}_i)}{\|\widetilde{v}_i\|\|\widetilde{w}_i\|} = \widetilde{\beta}_i > 0. \tag{4.11}$$

We denote the subspace correction method for the augmented system of equations based on the decomposition (4.10) and the bilinear forms $\widetilde{a}^i$ on $\widetilde{V}_i$ by (SSCA).

We now introduce the components needed for a subspace correction method (SSCO) for the original problem (1.1). For $V = \mathbb{R}^n$, we define $\Lambda : \widetilde{V} \mapsto V$ by

$$\Lambda\widetilde{v} = (\Phi, I)\widetilde{v} = (\Phi, I)\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \Phi v_1 + v_2. \tag{4.12}$$

It is obvious that the operator Λ is onto, but in general it may not be one-to-one. To obtain a space decomposition for V, each subspace V_i is chosen to be

$$V_i = \{v_i \in V : v_i = \Lambda\widetilde{v}_i,\ \forall \widetilde{v}_i \in \widetilde{V}_i\}.$$

Next, we introduce $a^i : V_i \times V_i \mapsto \mathbb{R}$, using the definition of $\widetilde{a}^i$ on $\widetilde{V}_i$.

$$a^i(v_i, w_i) = a^i(\Lambda\widetilde{v}_i, \Lambda\widetilde{w}_i) = \widetilde{a}^i(\widetilde{v}_i, \widetilde{w}_i),\ \forall \widetilde{v}_i, \widetilde{w}_i \in \widetilde{V}_i, \tag{4.13}$$

where $v_i = \Lambda\widetilde{v}_i$ and $w_i = \Lambda\widetilde{w}_i$. The mapping $\Lambda : \widetilde{V}_i \mapsto V_i$ is clearly one-to-one (on V_i) due to the solvability conditions (4.11) and therefore, it is an isomorphism. Furthermore, the bilinear forms $a^i(\cdot,\cdot)$ also satisfy the corresponding inf–sup conditions on V_i.

We now show that the two methods (SSCO) and (SSCA) are equivalent.

Theorem 4.1. *Let $\{u^\ell\}$ and $\{\widetilde{u}^\ell\}$ be two sequences of iterates generated by the methods (SSCO) and (SSCA) with initial guesses u^0 and $\widetilde{u}^0$, respectively. Then the algorithms (SSCA) and (SSCO) are equivalent in the following sense: If $u^0 = \Lambda\widetilde{u}^0$, then $u^\ell = \Lambda\widetilde{u}^\ell$.*

Proof. Let $\widetilde{u}^\ell$ denote the ℓth iterate obtained by (SSCA) and u^ℓ denote the ℓth iterate obtained by (SSCO). We will show that

$$u^{\ell+1} = \Lambda\widetilde{u}^{\ell+1}, \tag{4.14}$$

under the assumption that $\widetilde{u}^\ell$ and u^ℓ satisfy

$$u^\ell = \Lambda\widetilde{u}^\ell. \tag{4.15}$$

Thanks to the relation (3.2), we see that

$$(\widetilde{f}, \widetilde{v}_i) = \left(\begin{pmatrix} \Phi^t \\ I \end{pmatrix} f, \widetilde{v}_i \right) = (f, \Lambda \widetilde{v}_i), \quad \forall \widetilde{v}_i \in \widetilde{V}_i \tag{4.16}$$

and also

$$\widetilde{a}(\widetilde{v}, \widetilde{w}) = (\mathcal{A}\widetilde{v}, \widetilde{w}) = (A\Lambda\widetilde{v}, \Lambda\widetilde{w}) = a(\Lambda\widetilde{v}, \Lambda\widetilde{w}). \tag{4.17}$$

We consider now the local residual equations as follows:

$$a^i(e_i, v_i) = (f, v_i) - a(u_{i-1}^\ell, v_i), \qquad \forall\, v_i \in V_i,$$

$$\widetilde{a}^i(\widetilde{e}_i, \widetilde{v}_i) = (f, \Lambda\widetilde{v}_i) - \widetilde{a}(\widetilde{u}_{i-1}^\ell, \widetilde{v}_i) = (\widetilde{f}, \widetilde{v}_i) - \widetilde{a}(\widetilde{u}_{i-1}^\ell, \widetilde{v}_i), \quad \forall\, \widetilde{v}_i \in \widetilde{V}_i.$$

Therefore, If $u_{i-1}^\ell = \Lambda\widetilde{u}_{i-1}^\ell$, we get $e_i = \Lambda\widetilde{e}_i$ and $u_i^\ell = \Lambda\widetilde{u}_i^\ell$ directly by the above local residual equations. Since we assume that $u_0^\ell = u^\ell = \Lambda\widetilde{u}^\ell = \Lambda\widetilde{u}_0^\ell$ and from the aforementioned observation, we have $u_1^\ell = \Lambda\widetilde{u}_1^\ell$. Repeat this argument for $i = 2, \ldots, J$, we get $u^{\ell+1} = \Lambda\widetilde{u}^{\ell+1}$. Then, a simple induction shows that $u^\ell = \Lambda\widetilde{u}^\ell$ for all ℓ. □

As an illustration, let us revisit the example from Sec. 3, with dim $\mathcal{N}(A_0) = 1$ (see (3.5)). The Gauss–Seidel method for (3.5) is based on the decomposition

$$\widetilde{V} = \sum_{i=0}^{n} \widetilde{V}_i, \tag{4.18}$$

where $\widetilde{V}_i = \mathrm{span}\{\widetilde{e}_{i+1}\}$ for $i = 0, \ldots, n$. Here, $\widetilde{e}_i$ is the canonical basis for the space $\widetilde{V} = \mathbb{R}^{n+1}$. The equivalent subspace correction method for the original equation (3.4) is based on the space decomposition

$$V = \sum_{i=0}^{n} V_i, \tag{4.19}$$

where $V_i = \Lambda\widetilde{V}_i = \mathrm{span}\{e_i\}$ for $i = 1, \ldots, n$ and $V_0 = \mathrm{span}\{\xi\}$.

A simple conclusion that can be drawn from the aforementioned analysis is that if the null space is contained in one of subspaces, then the convergence of the subspace correction method will be ϵ-independent. However, in many cases it does not lead to an efficient solution method, because the dimension of the null space of A_0 may grow with the size of the problem. Nearly singular problems that exhibit such a behavior are easily found by considering discretizations of variational problems in $H(\mathrm{curl})$ or $H(\mathrm{div})$, because the dimension of the null space increases with the problem size and it is hard or even impossible to solve systems with $\Phi^T A/\Phi$ directly. Apparently, in such cases a good candidate for providing efficient iterative solver is a multigrid method. As it turns out, when designing a multigrid method, the role of the smoother (or equivalently the choice of the local subspaces) is crucial. As we have mentioned in Sec. 2, a point Gauss–Seidel relaxation will not work for discrete systems resulting from $H(\mathrm{curl})$ or $H(\mathrm{div})$ discretizations (see Ref. 37 for a proof of such result) and the local subspaces should contain the kernel functions to obtain a good smoother (see Ref. 2 and 3 for the relevant results on uniform convergence of multigrid method in such cases). In the next section, we formulate a general assumption that encompasses in a way these observations, and then prove a convergence result.

4.2. *Abstract assumptions*

In this section, we present the abstract assumptions under which we can obtain the parameter independent convergence of Algorithm 4.1. For the discussion of the parameter independent convergence, when inexact subspace solvers are used, we need some additional assumptions. We first assume that for each $i = 1, \dots, J$, there exists a constant $\delta_i \in [0, 1)$, independent on ϵ such that

$$\sup_{v_i \in V_i : \|v_i\|_a = 1} \|(I - T_i)v_i\|_a \leq \delta_i. \tag{4.20}$$

This assumption is equivalent to the following two assumptions: There exist $\omega_i \in (0, 2)$ and $\mu_i > 0$ independent of ϵ such that

$$(T_i v_i, T_i v_i)_a \leq \omega_i (T_i v_i, v_i)_a \quad \forall\, v_i \in V_i, \tag{4.21}$$

$$(T_i v_i, T_i v_i)_a \geq \mu_i (v_i, v_i)_a \qquad \forall\, v_i \in V_i. \tag{4.22}$$

We remark that the assumptions (4.21) and (4.22) are automatic for $T_i = P_i$ with $\omega_i = \mu_i = 1$. In particular, the assumption (4.21) is well known to be a necessary and sufficient condition for the energy norm convergence of the subspace correction method.[34,35] On the other hand, (4.22) is related to the limiting case $\varepsilon = 0$ and is necessary to guarantee ϵ-independent convergence. Similar assumptions have been introduced as necessary and sufficient conditions for the energy norm convergence of the semi-definite problems (case $\epsilon = 0$) in Refs. 26 and 25.

The next assumption is on the splitting of the null space of A_0. It plays a crucial role in the convergence analysis.

(A1) The decomposition $V = \sum_{i=1}^{J} V_i$ satisfies $\mathcal{N} = \sum_{k=1}^{J} (V_k \cap \mathcal{N})$.

In another word, **(A1)** implies that any element in $\mathcal{N}$ can be decomposed into a sum of elements in $V_k \cap \mathcal{N}$.

4.3. *Convergence estimate for Algorithm* 4.1

With all assumptions and notation in hand, we are ready to prove a convergence rate estimate for Algorithm 4.1, applied to problem (1.1).

The following lemmas are needed to prove the main results (Theorems 4.2 and 4.3).

Lemma 4.1. *The following inequality holds true.*

$$\|P_i u\|_a^2 \leq |P_{i,0} u|_{a_0}^2 + \epsilon \|P_{i,1} u\|_{a_1}^2, \quad \forall u \in V.$$

Proof. For any $u \in V$,

$$\begin{aligned} a(u, P_i u) &= a_0(u, P_i u) + \epsilon a_1(u, P_i u) \\ &= a_0(P_{i,0}u, P_i u) + \epsilon a_1(P_{i,1}u, P_i u) \\ &\leq \frac{1}{2}(a_0(P_i u, P_i u) + a_0(P_{i,0}u, P_{i,0}u)) \\ &\quad + \frac{\epsilon}{2}(a_1(P_i u, P_i u) + a_1(P_{i,1}u, P_{i,1}u)). \end{aligned}$$

This completes the proof. □

We now show the ϵ-independent convergence rate for the local subspace solver based on the assumptions (4.21) and (4.22).

Lemma 4.2. *Assume* (4.21) *and* (4.22). *Then we have*

$$\begin{aligned} (\bar{T}_i v_i, v_i)_a &= \|v_i\|_a^2 - \|(I - T_i)v_i\|_a^2 \\ &\geq \mu_i \left(\frac{2}{\omega_i} - 1\right) \|v_i\|_a^2, \qquad \forall v_i \in V_i, \end{aligned} \tag{4.23}$$

$$(T_i^* v_i, T_i^* v_i)_a \leq \omega_i^2 (v_i, v_i)_a, \qquad \forall v_i \in V_i, \tag{4.24}$$

and

$$(\bar{T}_i^{-1} v_i, v_i)_a \leq \frac{\omega_i}{\mu_i(2 - \omega_i)}(v_i, v_i)_a, \quad \forall v_i \in V_i. \tag{4.25}$$

Proof. The inequalities follow directly from the definitions and assumptions (4.21) and (4.22). □

The following convergence rate result for the subspace correction method for the symmetric and positive definite problems (see Ref. 26) will be crucial in our analysis.

Lemma 4.3. *If* (4.21) *holds, then*

$$\|E\|_a^2 = \sup_{\|v\|_a = 1} \|Ev\|_a^2 = 1 - \frac{1}{K},$$

with

$$K = \sup_{v \in V} \inf_{\sum_{i=1}^J v_i = v} \frac{\sum_{i=1}^J (\bar{T}_i^{-1} w_i, w_i)_a}{\|v\|_a^2},$$

where $w_i = v_i + T_i^* \sum_{j=i+1}^J v_j$.

Since an estimate on K also gives an estimate on the convergence rate $\|E\|_a$, below we only state the results in terms of K. Under the assumptions (4.21) and (4.22), we get an estimate which only depends on the exact subspace solvers and the constants ω_i and μ_i in (4.21) and (4.22).

Lemma 4.4. *Assume that* (4.21) *and* (4.22) *hold. Then,*

$$K \le \sup_{v\in V} \inf_{\sum_{i=1}^J v_i = v} \sum_{i=1}^{J} \frac{2\omega_i}{\mu_i(2-\omega_i)} \left(\|v_i\|_a^2 + \omega_i^2 \|P_i \sum_{j=i+1}^{J} v_j\|_a^2 \right) \Big/ \|v\|_a^2.$$

Proof. Set

$$w_i = v_i + T_i^* \sum_{j=i+1}^{J} v_j.$$

Using (4.21) and (4.22), and applying Lemma 4.2, leads to

$$\begin{aligned}
(\bar{T}_i^{-1} w_i, w_i)_a &\le \frac{\omega_i}{\mu_i(2-\omega_i)} \|w_i\|_a^2 \\
&\le \frac{\omega_i}{\mu_i(2-\omega_i)} \left\{ 2\|v_i\|_a^2 + 2 \left\| T_i^* P_i \sum_{j=i+1}^{J} v_j \right\|_a^2 \right\} \\
&\le \frac{2\omega_i}{\mu_i(2-\omega_i)} \left\{ \|v_i\|_a^2 + \omega_i^2 \left\| P_i \sum_{j=i+1}^{J} v_j \right\|_a^2 \right\}.
\end{aligned}$$

The proof is completed by applying Lemma 4.3. □

Lemma 4.5. *For any* $v \in V$, *we decompose* $v = w + \varphi$, *with* $w \in \mathcal{N}^\perp$ *and* $\varphi \in \mathcal{N}$, *then*

$$\|v\|_a^2 = |w|_{a_0}^2 + \epsilon\|w\|_{a_1}^2 + \epsilon\|\varphi\|_{a_1}^2. \tag{4.26}$$

Proof. Noting that $\mathcal{N}^\perp$ is the orthogonal complement of $\mathcal{N}$ with respect to $a_1(\cdot,\cdot)$ (see the beginning of Sec. 4), we have

$$\begin{aligned}
a(v,v) &= a(w+\varphi, w+\varphi) \\
&= a(w,w) + a(\varphi,\varphi) \\
&= a_0(w,w) + \epsilon a_1(w,w) + \epsilon a_1(\varphi,\varphi).
\end{aligned}$$

□

The first estimate that gives ϵ-independent convergence rate is for exact subspace solvers.

Theorem 4.2. *If* (**A1**) *holds, then Algorithm* 4.1 *with exact local subspace solves converges uniformly with respect to* ϵ *and we have the estimate*

$$\begin{aligned}
K \le{}& 2 \sup_{v\in\mathcal{N}^\perp} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \left(\frac{|P_{i,0}\sum_{j=i}^J v_j|_{a_0}^2}{|v|_{a_0}^2} + \frac{\|P_{i,1}\sum_{j=i}^J v_j\|_{a_1}^2}{\|v\|_{a_1}^2} \right) \\
&+ 2 \sup_{\varphi\in\mathcal{N}} \inf_{\sum_i \varphi_i = \varphi, \varphi_i\in\mathcal{N}} \frac{\sum_{i=1}^J \|P_{i,1}\sum_{j=i}^J \varphi_j\|_{a_1}^2}{\|\varphi\|_{a_1}^2}.
\end{aligned} \tag{4.27}$$

Proof. By Lemma 4.3, we get

$$K = \sup_{v \in V} \frac{K(v)}{\|v\|_a^2}, \quad \text{with } K(v) = \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \left\| P_i \sum_{j=i}^{J} v_j \right\|_a^2.$$

For a given $v \in V$, we have the decomposition $v = w + \varphi$ where $w \in \mathcal{N}^{\perp}$ and $\varphi \in \mathcal{N}$. By (**A1**), there exist $\varphi_i \in V_i \cap \mathcal{N}$ such that $\sum_i \varphi_i = \varphi$. Then

$$\begin{aligned} K(v) &\le \inf_{\sum_i v_i = w} \inf_{\sum_i \varphi_i = \varphi, \varphi_i \in \mathcal{N}} \sum_{i=1}^{J} \left\| P_i \sum_{j=i}^{J} (v_j + \varphi_j) \right\|_a^2 \\ &\le 2 \inf_{\sum_i v_i = w} \sum_{i=1}^{J} \left\| P_i \sum_{j=i}^{J} v_j \right\|_a^2 + 2 \inf_{\sum_i \varphi_i = \varphi, \varphi_i \in \mathcal{N}} \sum_{i=1}^{J} \left\| P_i \sum_{j=i}^{J} \varphi_j \right\|_a^2. \end{aligned}$$

By Lemma 4.5 and the above inequality, we get

$$\begin{aligned} K \le 2 & \sup_{v \in \mathcal{N}^{\perp}} \inf_{\sum_i v_i = v} \frac{\sum_{i=1}^{J} \|P_i \sum_{j=i}^{J} v_j\|_a^2}{|v|_{a_0}^2 + \epsilon \|v\|_{a_1}^2} \\ & + \frac{2}{\epsilon} \sup_{\varphi \in \mathcal{N}} \inf_{\sum_i \varphi_i = \varphi, \varphi_i \in \mathcal{N}} \frac{\sum_{i=1}^{J} \|P_i \sum_{j=i}^{J} \varphi_j\|_a^2}{\|\varphi\|_{a_1}^2}. \end{aligned}$$

By Lemma 4.1, we obtain that

$$\left\| P_i \sum_{j=i}^{J} v_j \right\|_a^2 \le \left| P_{i,0} \sum_{j=i}^{J} v_j \right|_{a_0}^2 + \epsilon \left\| P_{i,1} \sum_{j=i}^{J} v_j \right\|_{a_1}^2$$

and

$$\left\| P_i \sum_{j=i}^{J} \varphi_j \right\|_a^2 \le \left| P_{i,0} \sum_{j=i}^{J} \varphi_j \right|_{a_0}^2 + \epsilon \left\| P_{i,1} \sum_{j=i}^{J} \varphi_j \right\|_{a_1}^2 = \epsilon \left\| P_{i,1} \sum_{j=i}^{J} \varphi_j \right\|_{a_1}^2.$$

Combining the above three inequalities, we get the inequality (4.27). □

By Lemma 4.4 and using similar arguments as in the proof of Theorem 4.2, we get convergence result for the case of inexact subspace solvers.

Theorem 4.3. *Assume* (**A1**), (4.21) *and* (4.22) *hold true. Then the convergence rate of Algorithm* 4.1 *is given by* $\|E\|_a^2 = 1 - K^{-1}$ *with*

$$\begin{aligned} K \le 2 & \sup_{v \in \mathcal{N}^{\perp}} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \frac{2\omega_i}{\mu_i(2-\omega_i)} \left(\frac{|v_i|_{a_0}^2 + \omega_i^2 |P_{i,0} w_i|_{a_0}^2}{|v|_{a_0}^2} + \frac{\|v_i\|_{a_1}^0 + \omega_i^2 \|P_{i,1} w_i\|_{a_1}^2}{\|v\|_{a_1}^2} \right) \\ & + 2 \sup_{\varphi \in \mathcal{N}} \inf_{\sum_i \varphi_i = \varphi, \varphi_i \in \mathcal{N}} \sum_{i=1}^{J} \frac{2\omega_i}{\mu_i(2-\omega_i)} \frac{\|\varphi_i\|_{a_1}^2 + \omega_i^2 \|P_{i,1} w_{i,\varphi}\|_{a_1}^2}{\|\varphi\|_{a_1}^2}, \end{aligned}$$

where $w_i = \sum_{j=i+1}^{J} v_j$ *and* $w_{i,\varphi} = \sum_{j=i+1}^{J} \varphi_j$.

We remark that the estimate on K in Theorems 4.2 and 4.3 are also independent of the upper bound on ϵ. In the next section, we apply these estimates for discretizations of variational problems in $H(\text{grad})$, $H(\text{curl})$ and $H(\text{div})$.

5. H(grad), H(curl) and H(div) Systems

Let G represents any of the grad, curl or div operators. Given a bounded and convex polyhedral domain $\Omega \in \mathbb{R}^3$, we introduce the following Sobolev spaces:

$$H(G;\Omega) = \{v \in L^2(\Omega),\ Gv \in L^2(\Omega)\}.$$

We now consider the following model problems: Find $u \in H(G;\Omega)$ such that

$$a(u,v) = (Gu, Gv) + \epsilon(u,v) = (f,v), \quad v \in H(G;\Omega), \tag{5.1}$$

where $(\cdot,\cdot) = (\cdot,\cdot)_0$ is the L^2 inner product. (5.2) corresponds to the partial differential equation

$$G^*Gu + \epsilon u = f,$$

where G^* is the adjoint operator of G with respect to $(\cdot,\cdot)$ inner product, with natural (Neumann) boundary condition.

To introduce the finite element discretization for (5.1), we assume that the domain Ω is triangulated using simplexes and the triangulation is quasi-uniform. We denote this triangulation with $\mathcal{T}_h = \{\tau_h\}$. The spaces $V_h(G)$ are the conforming finite element spaces corresponding to the triangulation $\mathcal{T}_h$.

The finite element discretizations of (5.1) is: Find $u_h \in V_h(G)$ such that

$$a(u_h, v_h) = (f, v_h), \quad v \in V_h(G). \tag{5.2}$$

Examples for the lowest order $H(G,\Omega)$-conforming finite element spaces on a tetrahedral mesh $\mathcal{T}_h$ of Ω are listed below for various G.

$$\begin{aligned}
V_h(\text{grad}) &:= \{v_h \in H(\text{grad}) : v_{h|\tau}(\mathbf{x}) = a + \mathbf{b}\cdot\mathbf{x},\ a \in \mathbb{R}, \mathbf{b} \in \mathbb{R}^3,\ \forall\tau \in \mathcal{T}_h\},\\
V_h(\text{curl}) &:= \{\mathbf{v}_h \in H(\text{curl},\Omega) :\ \mathbf{v}_{h|\tau}(\mathbf{x}) = \mathbf{a} + \mathbf{x}\times\mathbf{b},\ \mathbf{a},\mathbf{b} \in \mathbb{R}^3,\ \forall\tau \in \mathcal{T}_h\},\\
V_h(\text{div}) &:= \{\mathbf{v}_h \in H(\text{div},\Omega) :\ \mathbf{v}_{h|\tau}(\mathbf{x}) = \mathbf{a} + \beta\mathbf{x},\ \mathbf{a} \in \mathbb{R}^3, \beta \in \mathbb{R}, \forall\tau \in \mathcal{T}_h\},\\
V_h(0) &:= \{v_h \in L^2(\Omega) :\ v_{h|\tau}(\mathbf{x}) = a,\ a \in \mathbb{R},\ \forall\tau \in \mathcal{T}_h\}.
\end{aligned}$$

The readers are refered to Refs. 23 and 28 for a detailed discussion on the finite element space $V_h(G)$. A simple description on $V_h(G)$ is made in Table 2.

Table 2. Finite element spaces of Whitney forms.

G	$H(G,\Omega)$	$V_h(G) \subset H(G,\Omega)$	FE space	Reference
grad	$H^1(\Omega)$	$V_h(\text{grad})$	Linear Lagrangian FE	18
curl	$H(\text{curl},\Omega)$	$V_h(\text{curl})$	Edge elements	28
div	$H(\text{div},\Omega)$	$V_h(\text{div})$	Face elements	28
0	$L^2(\Omega)$	$V_h(0)$	p.w. constants	

5.1. *Two-level method with Schwarz smoothers for* $\boldsymbol{H}$(curl) *and* $\boldsymbol{H}$(div) *systems*

To be specific, we will restrict our concern only on the two-level method with the exact coarse grid solver. We denote $\mathcal{T}_H = \{\tau_H\}$ by the triangulation of Ω with the mesh size H, and assume that $\mathcal{T}_h$ is obtained from the refinement of the triangulation $\mathcal{T}_H$. $V_H(G)$ denotes the conforming finite element spaces based on the triangulation $\mathcal{T}_H$.

Following the notation presented in our abstract convergence theories, we shall set $V = V_h(G)$ and

$$a(v, v) = a_0(u, v) + \epsilon(v, v),$$

where $(\cdot, \cdot) = (\cdot, \cdot)_0$ and $a_0(u, v) = (Gu, Gv)$. The null space of a_0 is denoted by $\mathcal{N}$ and V can be decomposed (orthogonal) as

$$V = \mathcal{N} \oplus \mathcal{N}^{\perp},$$

where $\perp$ is with respect to the L^2 inner product.

Now, we construct the subspace correction methods for (5.2) based on the vertex-based space decomposition as done in Ref. 3. For a vertex x^i of $\mathcal{T}_h$, we define

$$\mathcal{T}_h^i = \{\tau \in \mathcal{T}_h : x^i \in \tau\}, \quad \text{and} \quad \Omega_h^i = \text{interior}\left(\bigcup \mathcal{T}_h^i\right).$$

Then the domain Ω_h^i is the subdomain of Ω formed by the patch of elements with x^i as a vertex. The subspaces and the space decomposition are then given by

$$V_h^i = \{v \in V : \text{supp}(v) \subset \overline{\Omega_h^i}\}, \quad \text{and} \quad V = V_H + \sum_{i=1}^{J} V_h^i = \sum_{i=0}^{J} V_h^i, \tag{5.3}$$

where for convenience we have denoted $V_h^0 = V_H$ and J is the number of vertices of $\mathcal{T}_h$. We then apply the exact subspace solver in each subspace V_h^i as well as in the coarse space V_H. We call the aforementioned subspace correction methods as two-level method with vertex-based smoother. Observe that by the commuting diagram given in Fig. 1 (see also Ref. 3), we have

$$\mathcal{N} = \sum_{i=0}^{J} \mathcal{N}_i = \sum_{i=0}^{J} \mathcal{N} \cap V_h^i.$$

Thus, the decomposition (5.3) satisfies (**A1**). The following lemma is well-known (see Ref. 3).

Lemma 5.1. *Let* P_H *be the orthogonal projection onto* V_H *with respect to* $H(G)$ *inner product. Then under the assumption* $H \gtrsim h$ *and* $v \in (I - P_H)V_h$, *there exists a decomposition* $v = \sum_{i=1}^{J} v_i$ *with* $v_i \in V_h^i$ *such that*

$$\sum_{i=1}^{J} \|v_i\|_{H(G)}^2 \lesssim \|v\|_{H(G)}^2. \tag{5.4}$$

Furthermore, we can show the following result.

$$\begin{tikzcd}
R \arrow[r] & C^\infty \arrow[r, "\mathrm{grad}"] \arrow[d, "\Pi_h^{\mathrm{grad}}"] & C^\infty \arrow[r, "\mathrm{curl}"] \arrow[d, "\Pi_h^{\mathrm{curl}}"] & C^\infty \arrow[r, "\mathrm{div}"] \arrow[d, "\Pi_h^{\mathrm{div}}"] & C^\infty \arrow[r] \arrow[d, "\Pi_h^0"] & 0 \\
R \arrow[r] & V_h(\mathrm{grad}) \arrow[r, "\mathrm{grad}"] & V_h(\mathrm{curl}) \arrow[r, "\mathrm{curl}"] & V_h(\mathrm{div}) \arrow[r, "\mathrm{div}"] & L_h^2 \arrow[r] & 0
\end{tikzcd}$$

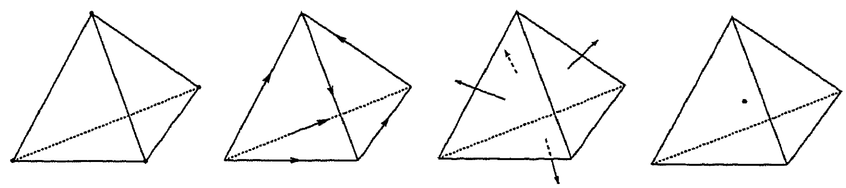

Fig. 1. Exact sequences, commutative diagrams, and degrees of freedom for the finite element spaces in the lowest order case.

Lemma 5.2. *Let $\Pi_H^G : V(G) \mapsto V_H(G)$ be the interpolation operator. Then under the assumption that $H \gtrsim h$ and $\varphi \in (I - \Pi_H^G)\mathcal{N}$, there exists a decomposition $\varphi = \sum_{i=1}^J \varphi_i$ with $\varphi_i \in \mathcal{N}_i$ such that*

$$\sum_{i=1}^J \|\varphi_i\|_0^2 \lesssim \|\varphi\|_0^2. \tag{5.5}$$

Proof. The proof for both curl and div systems is similar. Here, we only prove the case $G = \mathrm{curl}$. Then $\mathcal{N} = \{\mathrm{grad}\,\phi : \phi \in V_h(\mathrm{grad})\}$ and we have the following commuting relation (see Fig. 1),

$$\mathrm{grad}\,\Pi_H^{\mathrm{grad}} = \Pi_H^{\mathrm{curl}}\,\mathrm{grad}.$$

For $\varphi \in (I - \Pi_H^{\mathrm{curl}})\mathcal{N}$, there exists a function $\phi \in V_h(\mathrm{grad})$ such that

$$\varphi = \mathrm{grad}[I - \Pi_H^{\mathrm{grad}}]\phi.$$

Now, by the interpolating locally, there exists a decomposition of $\psi = (I - \Pi_H^{\mathrm{grad}})\phi = \sum_{i=1}^J \psi_i$ such that

$$\sum_{i=1}^J \|\psi^i\|_0^2 \lesssim \|\psi\|_0^2.$$

Then $\varphi = \sum_{i=1}^J \varphi_i = \sum_{i=1}^J \mathrm{grad}\,\psi_i$ is the desired decomposition, i.e.

$$\sum_{i=1}^J \|\mathrm{grad}\,\psi_i\|_0^2 \lesssim h^{-2} \sum_{i=1}^J \|\psi^i\|_0^2 \lesssim h^{-2}\|\psi\|_0^2 \lesssim \|\psi\|_1^2 = \|\varphi\|_0^2. \qquad \square$$

Note that in the setting described above, since $a_1(\cdot,\cdot) = (\cdot,\cdot)_0$, the projection operator $P_{i,1}$ is nothing else than the local L^2 projection for each $i = 1, \ldots, J$. Applying Theorems 4.2 and 4.3, one can obtain the following convergence rate estimate:

Lemma 5.3. *The two-level method with vertex-based Schwarz smoother for $H(\text{div})$ and $H(\text{curl})$ systems converges uniformly with respect to ϵ and the mesh size h.*

Proof. We have that $\|E\|_a^2 = 1 - \frac{1}{K}$ with

$$K \lesssim \sup_{v\in\mathcal{N}^\perp} \inf_{\sum_i v_i = v} \sum_{i=0}^{J} \left(\frac{|P_{i,0}\sum_{j=i}^J v_j|_{a_0}^2}{|v|_{a_0}^2} + \frac{\|P_{i,1}\sum_{j=i}^J v_j\|_0^2}{\|v\|_0^2} \right)$$
$$+ \sup_{\varphi\in\mathcal{N}} \inf_{\sum_i \varphi_i=\varphi, \varphi_i\in\mathcal{N}_i} \frac{\sum_{i=0}^J \|P_{i,1}\sum_{j=i}^J \varphi_j\|_0^2}{\|\varphi\|_0^2}. \tag{5.6}$$

For both systems, the last quantity can be similarly estimated by a generic constant, and here we shall estimate the first term only. We observe that by setting $v_0 = P_H v$,

$$\sum_{i=0}^J \left| P_{i,0} \sum_{j=i}^J v_j \right|_{a_0}^2 \lesssim \sum_{i=1}^J \left| P_{i,0}\sum_{j=i}^J v_j \right|_{a_0}^2 + \left| P_H v \right|_{a_0}^2$$
$$\lesssim \sum_{i=1}^J |v_j|_{a_0}^2 + |v|_{a_0}^2.$$

Similarly, we see that by setting $v_0 = P_H v$,

$$\sum_{i=0}^J \left\| P_{i,1} \sum_{j=i}^J v_j \right\|_0^2 \lesssim \sum_{i=1}^J \left\| P_{i,1}\sum_{j=i}^J v_j \right\|_0^2 + \|P_H v\|_0^2$$
$$\lesssim \sum_{i=1}^J \|v_j\|_0^2 + \|v\|_0^2.$$

With the choice of decompositions $\{v_j\}_{j=1}^J$ given in Lemma 5.1, we can conclude that

$$\sup_{v\in\mathcal{N}^\perp} \inf_{\sum_i v_i = v} \sum_{i=0}^{J} \left(\frac{|P_{i,0}\sum_{j=i}^J v_j|_{a_0}^2}{|v|_{a_0}^2} + \frac{\|P_{i,1}\sum_{j=i}^J v_j\|_0^2}{\|v\|_0^2} \right)$$
$$\lesssim \frac{|(I-P_H)v|_{a_0}^2 + |v|_{a_0}^2}{|v|_{a_0}^2} + \frac{\|(I-P_H)v\|_0^2 + \|v\|_0^2}{\|v\|_0^2}$$
$$\lesssim 1.$$

The above inequalities are obtained by the stability of the operator P_H. This completes the proof. □

5.2. *Multigrid analysis for $H(\text{grad}) = H^1(\Omega)$*

In this subsection, we will illustrate how our general theory may also be applied to multilevel analysis for the case $G = \text{grad}$. The technique presented in this section relies on the techniques introduced in Refs. 24 and 26 for semi-definite problems. The difference is that for $\epsilon > 0$, one can see the important role played by assumption (**A1**), in order to obtain $\{\epsilon, h\}$-independent convergence of Algorithm 4.1. We now consider the following problem in $H(\mathrm{grad})$:

$$-\Delta u + \epsilon u = f, \quad \text{in } \Omega, \tag{5.7}$$

subject to the pure Neumann boundary condition on $\partial\Omega$, i.e.

$$\mathbf{n} \cdot \nabla u = 0, \tag{5.8}$$

where $\mathbf{n}$ is the unit normal vector to $\partial\Omega$. The variational problem corresponding to (5.7) is as follows: Find $u \in H^1(\Omega)$ such that

$$a(u,v) = a_0(u,v) + \epsilon a_1(u,v) = (f,v), \quad \forall\, v \in H^1(\Omega), \tag{5.9}$$

where

$$a_0(u,v) = (u,v)_1 = \int_\Omega \nabla u \cdot \nabla v\, dx,$$

$$a_1(u,v) = (u,v)_0 = \int_\Omega uv\, dx,$$

and

$$(f,v) = \int_\Omega fv\, dx.$$

The null space $\mathcal{N}$ of $a_0(\cdot,\cdot)$ is $\mathcal{N} = \mathrm{span}\{1\}$. As usual, we assume that Ω is triangulated with a nested sequence of quasi-uniform triangles $\mathcal{T}_k = \{\tau_k^i\}$ of size h_k, where the quasi-uniformity constants are independent of k and $h_k \sim \gamma^k$ with $\gamma \in (0,1)$ for $k = 1, \ldots, J$. Associated with each $\mathcal{T}_k$, we have the finite element space of continuous piecewise linear functions $V_k \subset H^1(\Omega)$. In this setting, it is clear that

$$V_1 \subset \cdots \subset V_k \subset \cdots \subset V_J = V.$$

The standard nodal basis functions for each space V_k are denoted by ϕ_k^i, and we have the following decomposition of V:

$$V_k = \mathrm{span}\left\{\phi_k^1, \ldots, \phi_k^{n_k}\right\} = \sum_{i=1}^{n_k} V_k^i,$$

where $V_k^i = \mathrm{span}\{\phi_k^i\}$, $\dim V_k^i = 1$. We are interested in solving the system resulting from the standard finite element discretization: Find $u \in V$ with $h = h_J$ and $V = V_J$ such that

$$a(u,v) = (f,v), \quad \forall\, v \in V. \tag{5.10}$$

To solve problem (5.10), we consider a multigrid method with the Gauss–Seidel method as a smoother. Setting $V_0 = \mathcal{N}$, $n_0 = 1$, and $V_0^1 = V_0$, then V can be decomposed as follows:

$$V = \sum_{k=0}^{J} \sum_{i=1}^{n_k} V_k^i.$$

It is easy to see that the decomposition satisfies (**A1**). The error transfer operator, E, is

$$E = \Pi_{k=0}^{J}\Pi_{l=1}^{n_k}(I - P_k^l), \tag{5.11}$$

where P_k^l is the exact solver on V_k^l (see also Ref. 34). In the following discussion, $P_{k,0}^l$ and $P_{k,1}^l$ are the a_0 and a_1 orthogonal projections on V_k^l, respectively. We note that for all k and l, $P_{k,1}$ and $P_{k,1}^l$ are L^2 projections. Following more popular convention, we use the standard notation $Q_k = P_{k,1}$ and $Q_k^l = P_{k,1}^l$.

By a direct application of Theorem 4.2, we obtain the following convergence estimate.

Theorem 5.1.

$$\|E\|_a^2 < \delta < 1, \tag{5.12}$$

where δ is bounded uniformly with respect to the parameter ϵ, the number of levels J and the mesh size h.

Proof. By the Poincare's inequality and Theorem 4.2, we obtain

$$\begin{aligned}
K \lesssim 1 + \sup_{v\in\mathcal{N}^\perp} \inf_{\sum_{k=0}^J\sum_{i=1}^{n_k} v_k^i = v} &\left(\sum_{k=0}^{J}\sum_{i=1}^{n_k}\left|P_{k,0}^i \sum_{(\ell,j)\geq(k,i)} v_\ell^j\right|_1^2 \Bigg/ |v|_1^2\right.\\
&\left.+\sum_{k=0}^{J}\sum_{i=1}^{n_k}\left\|Q_k^i \sum_{(\ell,j)\geq(k,i)} v_\ell^j\right\|_0^2 \Bigg/ \|v\|_0^2\right)\\
&+\sup_{\varphi\in\mathcal{N}} \inf_{\sum_{k=0}^J\sum_{i=1}^{n_k}\varphi_k^i=\varphi,\varphi_k^i\in\mathcal{N}_k^i} \frac{\sum_{k=0}^J\sum_{i=1}^{n_k}\|Q_k^i\sum_{(\ell,j)\geq(k,i)}\varphi_\ell^j\|_0^2}{\|\varphi\|_0^2}.
\end{aligned}$$

For simplicity, we shall denote

$$\mathsf{I} = \sum_{k=0}^{J}\sum_{i=1}^{n_k}\left|P_{k,0}^i \sum_{(\ell,j)\geq(k,i)} v_\ell^j\right|_1^2 \Bigg/ |v|_1^2 + \sum_{k=0}^{J}\sum_{i=1}^{n_k}\left\|Q_k^i \sum_{(\ell,j)\geq(k,i)} v_\ell^j\right\|_0^2 \Bigg/ \|v\|_0^2,$$

$$\mathsf{II} = \sup_{\varphi\in\mathcal{N}} \inf_{\sum_{k=0}^J\sum_{i=1}^{n_k}\varphi_k^i=\varphi,\varphi_k^i\in\mathcal{N}_k^i} \sum_{k=0}^{J}\sum_{i=1}^{n_k}\left\|Q_k^i \sum_{(\ell,j)\geq(k,i)} \varphi_\ell^j\right\|_0^2 \Bigg/ \|\varphi\|_0^2.$$

Since V_0^1 is one dimensional, we have

$$\mathsf{II} = 2\sup_{\varphi\in\mathcal{N}} \frac{\|Q_0^1 c\|_0^2}{\|\varphi\|_0^2} = 2.$$

Now, by setting, $v_0^1 = 0$, we obtain the following estimate:

$$\mathrm{I} \lesssim 1 + \sup_{v \in \mathcal{N}^{\perp}} \inf_{\sum_{k=1}^{J} \sum_{i=1}^{n_k} v_k^i = v} \left(\sum_{k=1}^{J} \sum_{i=1}^{n_k} \left| P_{k,0}^i \sum_{(\ell,j) \geq (k,i)} v_\ell^j \right|_1^2 \Big/ |v|_1^2 \right.$$
$$\left. + \sum_{k=1}^{J} \sum_{i=1}^{n_k} \left\| Q_k^i \sum_{(\ell,j) \geq (k,i)} v_\ell^j \right\|_0^2 \Big/ \|v\|_0^2 \right).$$

The estimate of the right-hand side of the above inequality is standard, we refer to Ref. 26 for the details. □

6. Conclusions

We presented a transparent theoretical framework for proving convergence estimates for subspace correction algorithms for nearly singular problems. We have stated the minimal assumptions needed to provide parameter independent convergence. The abstract theory is applicable to a wide range of numerical models corresponding to finite element discretizations of partial differential equations, for example it can be used to design and analyze efficient methods for indefinite problems. As a future research, we plan to extend this work and derive convergence estimates for multigrid methods for anisotropic problems and linear elasticity problems.

Acknowledgments

The first author was supported in part by NSF DMS-0609655, the second author was supported by NSFC-10501001 and SRF for ROCS, SEM, the third author was supported in part by NSF DMS-0609727, DMS-0619587 and NSFC-10528102, and the fourth author was supported in part by NSF DMS-0511800 and DMS-0619587. All authors were also supported in part by the Center for Computational Mathematics and Applications in Pennsylvania State University.

References

1. A. B. Abdallah, F. B. Belgacem, Y. Maday and F. Rapetti, Mortaring the two-dimensional edge finite elements for the discretization of some electromagnetic models, *Math. Mod. Meth. Appl. Sci.* **14** (2004) 1635–1656.
2. D. N. Arnold, R. S. Falk and R. Winther, Preconditioning in $H(\mathrm{div})$ and applications, *Math. Comp.* **66** (1997) 957–984.
3. D. N. Arnold, R. S. Falk and R. Winther. Multigrid in $H(\mathrm{div})$ and $H(\mathrm{curl})$, *Numer. Math.* **85** (2000) 197–217.
4. O. Axelsson, *Iterative Solution Methods* (Cambridge Univ. Press, 1994).
5. I. Babuška, The finite element method with Lagrangian multipliers, *Numer. Math.* **20** (1973) 179–192.
6. J. W. Barrett, Christoph Schwab and Endre Süli, Existence of global weak solutions for some polymeric flow models, *Math. Mod. Meth. Appl. Sci.* **15** (2005) 939–983.

7. C. Bernardi, V. Girault and F. Hecht. *A posteriori* analysis of a penalty method and application to the Stokes problem, *Math. Mod. Meth. Appl. Sci.* **13** (2003) 1599–1628.
8. J. H. Bramble and X. Zhang, The analysis of multigrid methods, in *Handbook of Numerical Analysis*, Vol. VII (North-Holland, 2000).
9. S. C. Brenner and L. R. Scott, The mathematical theory of finite element methods, *Texts in Applied Mathematics*, Vol. 15, 2nd edn. (Springer-Verlag, 2002).
10. S. C. Brenner, A nonconforming multigrid method for the stationary Stokes equations, *Math. Comp.* **55** (1990) 411–437.
11. S. C. Brenner, A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity, *SIAM J. Numer. Anal.* **30** (1993) 116–135.
12. F. Brezzi, On the existence, uniqueness and approximation of saddle–point problems arising from Lagrange multipliers, *R.A.I.R.O Anal. Numer.* **R2** (1974) 129–151.
13. F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods* (Springer-Verlag, 1991).
14. Z. Q. Cai, C. I. Goldstein and J. E. Pasciak, Multilevel iteration for mixed finite element systems with penalty, *SIAM J. Sci. Comput.* **14** (1993) 1072–1088.
15. Z. Q. Cai, J. Mandel and S. McCormick, Multigrid methods for nearly singular linear equations and eigenvalue problems, *SIAM J. Numer. Anal.* **34** (1997) 178–200.
16. P. Ciarlet, *The Finite Element Method for Elliptic Problems* (North-Holland, 1978).
17. R. Falgout, P. Vassilevski and L. Zikatanov, On two-grid convergence estimates, *Numer. Linear Algebra Appl.*, **16** (2005) 471–494.
18. R. S. Falk, Nonconforming finite element methods for the equations of linear elasticity, *Math. Comp.* **57** (1991) 529–550.
19. M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems* (North-Holland, 1983).
20. Leopoldo P. Franca, Saulo P. Oliverira and Marcus Sarkis, Continuous Q1–Q1 stokes elements stabilized with nonconforming null edge average velocity functions, *Math. Mod. Meth. Appl. Sci.* **17** (2007) 439–459.
21. R. Glowinski, *Numerical Methods for Nonlinear Variational Problems* (Springer-Verlag, 1984).
22. M. Griebel, Multilevel algorithms considered as iterative methods on semidefinite systems, *SIAM J. Sci. Comput.* **15** (1994) 547–565.
23. R. Hiptmair, Finite elements in computational electromagnetism, *Acta Numer.* **11** (2002) 237–339.
24. Y.-J. Lee, Modelling and simulations of non-Newtonian fluid flows, PhD thesis, The Pennsylvania State University, University Park, PA (2004).
25. Y.-J. Lee, J. Wu, J. Xu and L. Zikatanov, On the convergence of iterative methods for semidefinite linear systems, *SIMAX* **28**(3) (2006) 634–641.
26. Y.-J. Lee, J. Wu, J. Xu and L. Zikatanov, A sharp convergence estimate of the method of subspace corrections for singular system of equations, to appear in *Math. Comp.*
27. Y.-J. Lee and J. Xu, New formulations, positivity preserving discretizations and stability analysis for non-Newtonian flow models, *Comput. Meth. Appl. Mech. Engrg.* **195** (2006) 1180–1206.
28. J. C. Nedelec, Mixed finite element in $\mathbb{R}^3$, *Numer. Math.* **35** (1980) 315–341.
29. A. Padiy, O. Axelsson and B. Polman, Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems, *SIAM J. Matrix Anal. Appl.* **22** (2000) 793–818 (electronic).
30. P.-A. Raviart and J. M. Thomas, Primal hybrid finite element methods for 2nd order elliptic problems, *Math. Comp.* **31** (1977) 391–413.

31. J. Schöberl, Multigrid methods for a parameter dependent problem in primal variables, *Numer. Math.* **84**(1) (1999) 97–119.
32. L. R. Scott and M. Vogelius, Conforming finite element methods for incompressible and nearly incompressible continua, *Large-Scale Computations in Fluid Mechanics, Part 2* (La Jolla, Calif., 1983), Lectures in Appl. Math., Vol. 22–2 (Amer. Math. Soc., 1985).
33. C. Wieners, Robust multigrid methods for nearly incompressible elasticity, *Computing* **64** (2000) 289–306. International GAMM-Workshop on Multigrid Methods (Bonn, 1998).
34. J. Xu, Iterative methods by space decomposition and subspace correction, *SIAM Rev.* **34** (1992) 581–613.
35. J. Xu and L. Zikatanov, The method of subspace corrections and the method of alternating projections in Hilbert space, *J. Amer. Math. Soc.* **15** (2002) 573–597.
36. Son-Young Yi, A new nonconforming mixed finite element method for linear elasticity, *Math. Mod. Meth. Appl. Sci.* **16** (2006) 979–999.
37. L. Zikatanov, Two sided bounds on the convergence rate of two-level methods, to appear in *Numer. Lin. Alg. Appl.*