THE METHOD OF ALTERNATING PROJECTIONS
AND THE METHOD OF SUBSPACE CORRECTIONS
IN HILBERT SPACE

JINCHAO XU AND LUDMIL ZIKATANOV

1. INTRODUCTION

The method of alternating projections and the method of subspace corrections are general iterative methods that have a variety of applications. The method of alternating projections, first proposed by von Neumann (1933) (see [31]), is an algorithm for finding the best approximation to any given point in a Hilbert space from the intersection of a finite number of subspaces. The method of subspace corrections, an abstraction of general linear iterative methods such as multigrid and domain decomposition methods, is an algorithm for finding the solution of a linear system of equations. In this paper, we shall study these two methods in a Hilbert space setting and in particular present a new identity for the product of nonexpansive operators that gives a sharpest possible estimate of the convergence rate of these methods.

Let V be a Hilbert space and $V_i \subset V$ $(i = 1, \ldots, J)$ a number of closed subspaces satisfying $V = \sum_{i=1}^{J} V_i$. One main result in this paper is that the following identity holds for an appropriate class of operators $T_i : V \mapsto V_i$ (see Theorem 4.2 below):

(1.1)
$$\|(I - T_J) \cdots (I - T_1)\|_{\mathcal{L}(V,V)}^2 = 1 - \left(1 + \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} (\bar{T}_i^{-1} T_i^* w_i, T_i^* w_i) \right)^{-1},$$

with $w_i = \sum_{j=i}^{J} v_j - T_i^{-1} v_i$ and $\bar{T}_i = T_i^* + T_i - T_i^* T_i$.

Roughly speaking, the above identity holds if each $I - T_i$ is nonexpansive. For example, if each $T_i = P_i : V \mapsto V_i$ is the orthogonal projection with respect to the underlying inner product of V, the above identity reads:

(1.2)
$$\|(I - P_J) \cdots (I - P_1)\|_{\mathcal{L}(V,V)}^2 = 1 - \left(1 + \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \|P_i \sum_{j=i+1}^{J} v_j\|^2 \right)^{-1}.$$

These identities give a precise quantitative estimate of the convergence rate of the method of alternating projections and the method of subspace corrections. Such identities are very desirable from the viewpoints of both theory and application. The sup-inf term in these identities is easy to estimate in applications and many existing convergent results can be derived from it in a straightforward fashion. It is expected that new convergence estimates can also result from such an identity for various applications. This result is also expected to be instrumental in the design of the method of subspace corrections or alternating projections because it is an *identity*, from which we know better than from any earlier theory what contributes to the rate of convergence and hence know where to look for improvement.

The method of subspace corrections has often been discussed in finite dimensional spaces in the literature since practical applications of this method are often for problems posed in a discrete setting in finite dimensions. In this paper, however, we shall discuss this method in a general Hilbert space setting. From a practical point of view, this level of abstraction may not be essential, but working on a general Hilbert space (which may be of infinite dimension) seems to make us see things in a more appropriate way.

In addition to the aforementioned identities, we shall also present a number of other important results in relation to these two types of methods. For example, we shall prove that the method of alternating projections is in fact equivalent to the method of subspace corrections in certain circumstances.

The rest of the paper is organized as follows. In §2 we introduce a framework of the method of subspace corrections. In §3 we consider the method of alternating projections and prove that it is equivalent to a special case of the method of subspace corrections. The main result of the paper is contained in §4 (Theorem 4.2) and §5 illustrates how our new identity can be used to obtain estimates of the convergence rate in multigrid and domain decomposition methods.

2. MSC: The method of subspace corrections

We shall present here the method of subspace corrections in the framework of solving a linear variational problem. Let H be a Hilbert space and $V \subset H$ a closed subspace. We are interested in solving the following variational problem: Find $u \in V$ for any given $f \in H^*$ such that

$$(2.1) \qquad a(u, v) = \langle f, v \rangle, \quad \forall v \in V.$$

Here $a(\cdot, \cdot) : H \times H \mapsto \mathbb{R}$ is a continuous bilinear form satisfying

$$a(u, v) \le \|a\| \|u\| \|v\|, \quad u, v \in H.$$

We assume that it further satisfies the following conditions:

$$(2.2) \qquad \inf_{u \in V} \sup_{v \in V} \frac{a(u, v)}{\|u\| \|v\|} = \inf_{v \in V} \sup_{u \in V} \frac{a(u, v)}{\|u\| \|v\|} > 0.$$

These conditions are often known as inf-sup conditions or Babuška-Brezzi conditions. It is known that these are the necessary and sufficient conditions for the well-posedness of the variational problem (2.1) (see [2], [14], and also [37]).

We would like to remark that the problem (2.1) often occurs with $V = H$ in most applications. The added twist here of using two possibly different spaces V and H makes it more convenient to discuss the relationship between the method of

subspace corrections and the method of alternating projections (see Theorem 3.3 below).

We shall now discuss an iterative procedure for solving (2.1). Given an initial guess u^0, an iterative method produces a sequence of u^ℓ ($\ell = 0, 1, 2, \ldots$) that (hopefully) better and better approximate the exact solution u. A typical construction of a linear iterative procedure can be described as follows. Assume $u^{\ell-1}$ is given; then we can define $u^\ell = u^{\ell-1} + \hat{e}$ where $\hat{e}$ is an approximate solution of the following residual equation:

$$(2.3) \qquad a(e, v) = f(v) - a(u^{\ell-1}, v), \quad \forall v \in V.$$

The residual equation is in general as difficult to solve as the original problem, but it is possible to solve it approximately (economically) to produce certain corrections on $u^{\ell-1}$ to obtain a better approximation u^ℓ.

The idea of the method of subspace corrections is to solve the residual equation (2.3) on some properly chosen subspaces.

We consider a collection of closed subspaces

$$V_i \subset V, \quad i = 1, 2, \ldots, J,$$

such that V is the closure of $\sum_i V_i$, namely

$$(2.4) \qquad V = \overline{\sum_{i=1}^{J} V_i}.$$

Associated with each subspace V_i, we introduce a bilinear form $a_i(\cdot, \cdot)$ which can be viewed as an approximation of $a(\cdot, \cdot)$ restricted on V_i. To assure the well-posedness of the subspace problems, we assume that the following inf-sup conditions are satisfied for all $i = 1, 2, \ldots, J$:

$$(2.5) \qquad \inf_{u_i \in V_i} \sup_{v_i \in V_i} \frac{a(u_i, v_i)}{\|u_i\| \|v_i\|} = \inf_{v_i \in V_i} \sup_{u_i \in V_i} \frac{a(u_i, v_i)}{\|u_i\| \|v_i\|} > 0$$

and

$$(2.6) \qquad \inf_{u_i \in V_i} \sup_{v_i \in V_i} \frac{a_i(u_i, v_i)}{\|u_i\| \|v_i\|} = \inf_{v_i \in V_i} \sup_{u_i \in V_i} \frac{a_i(u_i, v_i)}{\|u_i\| \|v_i\|} > 0.$$

2.1. **SSC: Successive subspace corrections.** The *method of successive subspace corrections* (MSSC) is an iterative algorithm that corrects residual equation successively on each subspace.

Algorithm 2.1 (MSSC). Let $u^0 \in H$ be given.

> **for** $\ell = 1, 2, \ldots$
> $\qquad u_0^{\ell-1} = u^{\ell-1}$
> $\qquad$ **for** $i = 1 : J$
> $\qquad\qquad$ Let $e_i \in V_i$ solve
>
> $$(2.7) \qquad a_i(e_i, v_i) = f(v_i) - a(u_{i-1}^{\ell-1}, v_i) \quad \forall v_i \in V_i$$
>
> $\qquad\qquad u_i^{\ell-1} = u_{i-1}^{\ell-1} + e_i$
> $\qquad$ **endfor**
> $\qquad u^\ell = u_J^{\ell-1}$
> **endfor**

We note that the above algorithm is well defined, thanks to the assumptions (2.5) and (2.6). For the analysis of this algorithm, let us introduce another class of linear operators $T_i : V \mapsto V_i$ that are defined by:

$$(2.8) \qquad a_i(T_i v, v_i) = a(v, v_i), \quad \forall v_i \in V_i.$$

Again, because of (2.5) and (2.6), each T_i is well defined and

$$(2.9) \qquad \mathcal{R}(T_i) = V_i \text{ and } T_i : V_i \mapsto V_i \text{ is isomorphic.}$$

In the special case when the subspace equation is solved exactly, we shall use the notation P_i for T_i, namely

$$P_i = T_i \text{ if } a_i(\cdot, \cdot) = a(\cdot, \cdot),$$

or P_i is defined by the following relation:

$$a(P_i v, v_i) = a(v, v_i), \quad v \in V, v_i \in V_i.$$

P_i is idempotent, namely $P_i^2 = P_i$. It is an orthogonal projection when $a(\cdot, \cdot)$ is an inner product itself, namely when $a(\cdot, \cdot)$ is symmetric positive definite. We note that the following relation holds:

$$(2.10) \qquad T_i = T_i P_i.$$

It is easy to see that

$$u - u_i^{\ell-1} = (I - T_i)(u - u_{i-1}^{\ell-1}).$$

A recursive application of the above identity yields

$$(2.11) \qquad u - u^\ell = E(u - u^{\ell-1}) = \cdots = E^\ell(u - u^0)$$

where

$$(2.12) \qquad E = (I - T_J)(I - T_{J-1}) \cdots (I - T_1).$$

Because of this special form of E, the error propagation operator, the successive subspace corrections method is also known as the multiplicative or product (Schwarz) method.

The general notion of subspace corrections by means of space decomposition (2.4) was described in Xu [35] based on the works of Bramble, Pasciak, Wang and Xu [10, 9]. It is an abstract generalization of a large class of iterative algorithms mostly used for solving systems of equations arising from the discretization of partial differential equations, such as Gauss-Seidel, multigrid and domain decomposition methods. In the last two decades a lot of effort has been put into the investigation of the theoretical and practical issues related to these methods. For a literature review and basic results we refer the reader to some monographs and survey articles: Hackbusch [25, 24], Xu [34, 35, 36], Xu and Zou [38], Yserentant [39], Bramble [8], and more recent works on this subject: Bramble and Zhang [13], Trottenberg, Oosterlee and Schüller [30], Smith, Bjøstrad and Gropp [29].

2.2. **Parallel subspace corrections.** Related to the method of successive subspace corrections discussed above, the method of parallel subspace corrections (see [35]) is another important class of algorithms. In this algorithm, the residual corrections are done in parallel in each subspace. The parallel nature of the algorithm makes it attractive for parallel computations.

A typical parallel subspace corrections method may be described as follows.

Algorithm 2.2 (MPSC)**.** Let $u^0 \in H$ be given.

> **for** $\ell = 1, 2, \ldots$
> > **for** $i = 1 : J$
> > > Let $e_i \in V_i$ solve
> > > $$a_i(e_i, v_i) = f(v_i) - a(u^{l-1}, v_i) \quad \forall v_i \in V_i$$
> > **endfor**
> > $u^\ell = u^{\ell-1} + \sum_{i=1}^{J} e_i$
> **endfor**

It is easy to see that $u - u^\ell = (I - T)(u - u^{\ell-1})$, where

$$(2.13) \qquad\qquad T = \sum_{i=1}^{J} T_i.$$

One more robust way of using parallel subspace correction to solve for (2.1) is to use the operator T and solve the following equivalent equation:

$$(2.14) \qquad\qquad Tu = w$$

where $w = \sum_{i=1}^{J} w_i$ with $w_i \in V_i$ obtained by solving

$$a_i(w_i, v_i) = \langle f, v_i \rangle, \quad \forall v_i \in V_i.$$

The system (2.14) is often known as a preconditioned system of the original system (2.1). Most often we expect that the operator T is well conditioned, namely it has a relatively small condition number:

$$\mathrm{cond}(T) \equiv \|T\| \, \|T^{-1}\|.$$

Because of the special form of T in (2.13), the method of parallel subspace corrections is also known as the additive (Schwarz) method. In the context of the multigrid method, (2.13) is closely related to the BPX-preconditioner (see Bramble, Pasciak and Xu [11]).

For the study of the operator T, let us make the following assumption:

$$(\mathbf{A0}) \qquad \sum_{i=1}^{J} V_i \text{ is closed, namely } V = \sum_{i=1}^{J} V_i.$$

which, by a simple application of the Open Mapping Theorem, implies

$$(2.15) \qquad\qquad \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \|v_i\|^2 < \infty.$$

The role of ($\mathbf{A0}$) can be seen in the following simple result.

Theorem 2.3. *The following two statements are equivalent:*

(1) $\sum_{i=1}^{J} V_i$ *is closed, namely* ($\mathbf{A0}$) *is satisfied.*

(2) $T = \sum_{i=1}^{J} T_i$ *is an isomorphism on V if each $T_i : V \mapsto V_i$ is a symmetric positive definite isomorphism on V_i.*

Proof. We first note that (2) obviously implies (1) since each $v \in V$ can be written as $v = \sum_i T_i T^{-1} v$.

Now we assume that (1) holds. We consider the following norm for $\widetilde{V} \equiv V_1 \times V_2 \times \ldots \times V_J$:

$$\|\tilde{v}\|^2 = \sum_{i=1}^{J} (T_i^{-1} v_i, v_i).$$

We define, as in (4.17),

$$\underline{T} : \widetilde{V} \mapsto V, \quad \underline{T}\tilde{v} = \sum_{i=1}^{J} T_i v_i.$$

Note that $\mathcal{R}(\underline{T}) = \sum_{i=1}^{J} V_i$ since each T_i is an isomorphism on V_i. By (1), $\mathcal{R}(\underline{T}) = V$ which implies that there exists a constant $\kappa > 0$ such that

$$(Tv, v) = \|\underline{T}^* v\|^2 \geq \kappa \|v\|^2, \quad \forall v \in V.$$

This implies (2). □

We shall now include one simple but important result on the operator T. Related results can be found in [35, 33, 22].

Lemma 2.4. *If* (**A0**) *is satisfied and each T_i is symmetric positive definite on V_i, then*

$$(2.16) \qquad (T^{-1}v, v) = \inf_{\sum v_i = v} \sum_{i=1}^{J} (T_i^{-1} v_i, v_i).$$

Proof. To prove (2.16) we take $v \in V$ and consider $v_i = T_i T^{-1} v$. Obviously $v = \sum_{i=1}^{J} v_i$. Then we have

$$
\begin{aligned}
\inf_{\sum u_i = v} \sum_{i=1}^{J} (T_i^{-1} u_i, u_i) &= \inf_{\sum w_i = 0} \sum_{i=1}^{J} (T_i^{-1}(v_i + w_i), v_i + w_i) \\
&= (T^{-1}v, v) + \inf_{\sum w_i = 0} \left[\sum_{i=1}^{J} 2(Tv, w_i) + (T_i^{-1} w_i, w_i) \right] \\
&= (T^{-1}v, v).
\end{aligned}
$$

□

The above result can be used to estimate the condition number of T. For example, the following estimate is very useful:

$$(2.17) \qquad (\lambda_{\min}(T))^{-1} = \sup_{\|v\|=1} \inf_{\sum v_i = v} \sum_{i=1}^{J} (T_i^{-1} v_i, v_i).$$

3. MAP: The method of alternating projections

In this section, we shall discuss another class of iterative methods, namely the method of alternating projections, and its relationship with the method of subspace corrections.

3.1. **A description of the method of alternating projections.** The method of alternating projections is, in its simplest form, due to von Neumann [31]. Let us now briefly describe this method. Let H be a Hilbert space and let $M_1, M_2 \subset H$ be two closed subspaces. Let P_{M_1} and P_{M_2} be two orthogonal projections from H to M_1 and M_2 respectively. It is easy to see that $P_{M_1} P_{M_2} = P_{M_1 \cap M_2}$ if (and only if) P_{M_1} and P_{M_2} commute, namely $P_{M_1} P_{M_2} = P_{M_2} P_{M_1}$. Von Neumann [31] proved that, even if P_{M_1} and P_{M_2} do not commute, the following identity holds:

$$\lim_{k \to \infty} (P_{M_2} P_{M_1})^k = P_{M_1 \cap M_2}.$$

The above result generalizes to more than two subspaces in a straightforward fashion. Let $M_i \subset H$ be closed subspaces or affine subspaces for $i = 1 : J$. Let P_{M_i} be orthogonal projections from H to M_i. Given $w \in H$, the MAP algorithm for finding the projection $P_M w$ where $M = \bigcap_{i=1}^J M_i$ is as follows.

Algorithm 3.1 (MAP). **Set** $w^0 = w$.

> **for** $\ell = 1, 2, \dots$
> $\quad w_0^{\ell-1} = w^{\ell-1}$
> $\quad$ **for** $i = 1 : J$
> $\quad\quad w_i^{\ell-1} = P_{M_i} w_{i-1}^{\ell-1}$
> $\quad$ **endfor**
> $\quad w^\ell = w_J^{\ell-1}$
> **endfor**

Roughly speaking, the iterates $\{w^\ell\}$ are obtained by alternatively computing the projections onto individual subspaces M_i. This method is thus useful when each P_{M_i} is "easier" to compute than P_M. We note that

$$(3.1) \qquad w^\ell = (P_{M_J} P_{M_{J-1}} \cdots P_{M_1}) w^{\ell-1} = (P_{M_J} P_{M_{J-1}} \cdots P_{M_1})^\ell w.$$

3.2. **Relationship between MSC and MAP.** The method of alternating projections is closely related to the method of subspace corrections. We shall now discuss their relationship.

Let us first include (and prove for completeness) the following well-known result.

Lemma 3.2. *Let M_i be closed subspaces of H and $M = \bigcap_{i=1}^J M_i$. Then*

$$M^\perp = \overline{\sum_{i=1}^J M_i^\perp}.$$

Proof. Denote $V = \sum_{i=1}^J M_i^\perp$. Let $u = \sum_i u_i \in V$. For each i, $(u_i, v) = 0$ for all $v \in M \subset M_i$ and hence $(u, v) = \sum_i (u_i, v) = 0$. This proves that $V \subset M^\perp$. Hence $\bar{V} \subset M^\perp$ since $M^\perp$ is closed.

On the other hand, if $u \in V^\perp$, then $(u, v_i) = 0$ for each i and $v_i \in M_i^\perp \subset V$. This means that $u \in (M_i^\perp)^\perp = M_i$ for each i and hence $u \in M$. This proves that $V^\perp \subset M$ and hence $M^\perp \subset (V^\perp)^\perp = \bar{V}$. $\qquad\qquad\square$

Theorem 3.3. *Let $M_i \subset H$ be closed subspaces and $M = \bigcap_{i=1}^J M_i$. Given $w \in H$, Algorithm 3.1 (MAP) for computing P_M by using P_{M_i} is equivalent to Algorithm 2.1 (MSC) for solving (2.1) if*

(1) $a(\cdot, \cdot) = (\cdot, \cdot)_H$;
(2) $f \in H^*$ is given by $\langle f, \phi \rangle = (w, \phi)_H, \quad \forall \phi \in H$;

(3) $V = \sum_{i=1}^{\overline{J}} V_i$ with $V_i = M_i^{\perp}$;

(4) $a_i(\cdot, \cdot) = a(\cdot, \cdot)$, namely $T_i = P_i$ for each i;

(5) $u^0 = 0$.

Furthermore

$$u = P_V w = P_M^{\perp} w, u^{\ell} = w - w^{\ell}.$$

Proof. For Algorithm 2.1, with (1)–(5), we have, by (2.11)

$$u - u^{\ell} = [(I - P_{V_J}) \cdots (I - P_{V_1})]^{\ell} u = [P_{M_J} \cdots P_{M_1}]^{\ell} u.$$

Thanks to Lemma 3.2, $V = M^{\perp}$. Hence $u = P_V w = w - P_M w$. Thus

$$w - P_M w - u^{\ell} = [P_{M_J} \cdots P_{M_1}]^{\ell}(w - P_M w) = [P_{M_J} \cdots P_{M_1}]^{\ell} w - P_M w.$$

Namely

$$w - u^{\ell} = [P_{M_J} \cdots P_{M_1}]^{\ell}(w - P_M w) = [P_{M_J} \cdots P_{M_1}]^{\ell} w$$

which, by (3.1), is w^{ℓ} from Algorithm 3.1. This completes the proof. $\qquad\square$

The relationship between the method of alternating projections and the method of subspace corrections has been observed by some authors although their algorithmic equivalence has not been explicitly stated before in the literature. For example, in Gilbert and Light [21], it was observed that the multigrid method (which is a method of subspace corrections; see [35] and §5) can be analyzed in the theoretical framework of alternating projections (although it was also stated there that the multigrid method, as an *algorithm*, would *not* be a method of alternating projections).

3.3. **On the convergence of MAP.** It has been proved (von Neumann [31] for $J = 2$ and Halperin [26] for $J \geq 2$) that

$$\lim_{\ell \to \infty} \|(P_{M_J} P_{M_{J-1}} \cdots P_{M_1})^{\ell} w - P_M w\| = 0 \quad \text{for all } w \in H.$$

The following rate of convergence is known (see Aronszajn [1] and Kayalar and Weinert [27]):

$$\|(P_{M_1} P_{M_2})^k - P_{M_1 \cap M_2}\| = c^{2k-1}(M_1, M_2)$$

where $c(M_1, M_2)$ is the cosine of the angle between M_1 and M_2:

$$c(M_1, M_2) = \sup\{\frac{(u, v)}{\|u\|\|v\|} : u \in M_1 \cap (M_1 \cap M_2)^{\perp}, \quad v \in M_2 \cap (M_1 \cap M_2)^{\perp}\}.$$

Such an estimate has been rediscovered several times, and we refer to Deutsch [15, 16, 17] for a more detailed discussion and literature review. An estimate like this has been tried to generalize to the case of more than two subspaces and similar (but less sharp) estimates for the rate of convergence have also been obtained in the literature (see Deutsch [18], Bauschke and Borwein [5], Bauschke, Deutsch and Hundel [4]).

It is interesting to note that, despite the close relationship between MAP and MSC, the aforementioned estimates for MAP in the literature are quite different in nature from those for MSC. In the next section, we shall present an identity for the product of operators (such as projections) and this identity can then be used for sharp convergence estimates for both the method of subspace corrections and

the method of alternating projections. Our new convergence theory is more closely related to those for the method of subspace corrections (cf. [35, 13]).

4. A new identity for the product of nonexpansive operators

The aim of this section is to present a new identity for the norm of a product of nonexpansive operators. This identity can be used for estimating the convergence rate of the method of subspace corrections and the method of alternating projections. We shall first confine our discussion on the method of subspace corrections and use the relevant notation.

Given a Hilbert space H and a closed subspace $V \subset H$, the method of subspace corrections is based on a collection of closed subspaces $V_i \subset V$, $i = 1, 2, \ldots, J$, satisfying

$$(4.1) \qquad V = \overline{\sum_{i=1}^{J} V_i}.$$

In view of (2.11), the convergence of the method of subspace corrections is equivalent to

$$(4.2) \qquad \lim_{\ell \to \infty} E^\ell = 0$$

where E is given by (2.12), namely

$$E = (I - T_J)(I - T_{J-1}) \cdots (I - T_1).$$

We shall establish such a convergence result under some natural conditions on subspace operators T_i. In fact, we shall establish a uniform convergence result by proving that

$$(4.3) \qquad \|E\| < 1$$

under certain assumptions.

There are a lot of works devoted to the estimate of $\|E\|$ (see [9, 35, 13]). But the estimate that we will present is the sharpest possible. In fact, we shall give an identity for $\|E\|$, rather than upper bound estimates as in all other works.

4.1. Assumptions on subspace solvers. We now try to derive conditions on the subspaces and subspace solvers for the convergence of the MSC.

First of all, we assume that

$$(\mathbf{A1}) \qquad \mathcal{R}(T_i) = V_i \text{ and } T_i : V_i \mapsto V_i \text{ is isomorphic for each } i = 1 : J.$$

If T_i is given by (2.8), as mentioned earlier, this assumption is a consequence of the inf-sup conditions (2.5) and (2.6).

In order for (4.2) to be valid, it is natural to expect that each factor $I - T_i$ should be nonexpansive, namely

$$(4.4) \qquad \|I - T_i\| \leq 1$$

which is equivalent to

$$(4.5) \qquad (T_i v, T_i v) \leq 2(T_i v, v), \quad v \in V.$$

There are operators T_i satisfying the above inequality but which may not lead to (4.2). One simple example is that $T_i = 2I$. In this case, $I - T_i = -I$ and (4.5) is satisfied with equality. To avoid such type of operators, we are led to make the following assumption on each T_i:

(**A2**) $\|T_i v\|^2 \leq \omega(T_i v, v), \quad v \in V$ for some constant $\omega \in (0, 2)$.

In the special case that $a(\cdot, \cdot) = (\cdot, \cdot)$, the above condition is equivalent to the same condition only in the subspace:

$$(4.6) \qquad \|T_i v_i\|^2 \leq \omega(T_i v_i, v_i), \quad v_i \in V_i.$$

The assumption (4.6) has been used in other papers; see [13] and the references cited therein. As we shall see below, this condition means that $I - T_i$ is a contraction on the subspace V_i.

Associated with each T_i, we introduce its symmetrization (see [35]):

$$(4.7) \qquad \bar{T}_i = T_i + T_i^* - T_i^* T_i.$$

This operator will play an important role in our analysis.

Lemma 4.1. *Assume that T_i satisfies (**A1**) and (**A2**). Then*

(1) *$I - T_i$ is nonexpansive.*
(2) *T_i, T_i^* and $\bar{T}_i$ have the same kernel: $\mathcal{N}(\bar{T}_i) = \mathcal{N}(T_i) = \mathcal{N}(T_i^*)$.*
(3) *T_i, T_i^* and $\bar{T}_i$ have the same range: $\mathcal{R}(\bar{T}_i) = \mathcal{R}(T_i) = \mathcal{R}(T_i^*) = V_i$.*
(4) *The following inequality holds:*

$$\frac{2-\omega}{\omega}\|T_i v\|^2 \leq \|v\|^2 - \|(I - T_i)v\|^2 = (\bar{T}_i v, v), \quad v \in V.$$

(5) *As operators restricted on V_i, the above (1)–(3) are still valid.*
(6) *T_i, T_i^* and $\bar{T}_i$ are all isomorphisms from V_i to itself.*
(7) *$\bar{T}_i$ is nonnegative on V and symmetric positive definite on V_i.*

Proof. It is obvious that (1) holds. By (1) $\|I - T_i\| \leq 1$, namely $\|T_i v\|^2 \leq 2(T_i v, v) = 2(v, T_i^* v)$; we see immediately that $\mathcal{N}(T_i^*) \subset \mathcal{N}(T_i)$. Similarly $\mathcal{N}(T_i^*) \subset \mathcal{N}(T_i)$ since $\|I - T_i^*\| = \|I - T_i\|$. Hence $\mathcal{N}(T_i) = \mathcal{N}(T_i^*)$. It follows from the definition of $\bar{T}_i$ that $\mathcal{N}(T_i) \subset \mathcal{N}(\bar{T}_i)$. Note that (**A2**)

$$(4.8) \qquad (\bar{T}_i v, v) = \|v\|^2 - \|(I - T_i)v\|^2 = 2(T_i v, v) - \|T_i v\|^2 \geq \frac{2-\omega}{\omega}\|T_i v\|^2$$

which implies $\mathcal{N}(T_i) \supset \mathcal{N}(\bar{T}_i)$. This completes the proof of (2). (3) follows from (2) by the well-know relations between kernel and range of operators. (4) is already contained in (4.8). (5)–(7) are obvious. $\qquad\square$

4.2. **Main result.** We are now in a position to present the main result of this paper. The theorem presented below is based on the aforementioned (**A1**), (**A2**) and also (**A0**) introduced in §2.2, namely

$$(\textbf{A0}) \qquad \sum_{i=1}^{J} V_i \text{ is closed, namely } V = \sum_{i=1}^{J} V_i.$$

Theorem 4.2. *Under the assumptions (**A0**), (**A1**) and (**A2**), the following identity holds:*

$$(4.9) \qquad \|E\|_{\mathcal{L}(V,V)}^2 \equiv \|(I - T_J)(I - T_{J-1}) \cdots (I - T_1)\|^2 = \frac{c_0}{1 + c_0}$$

where

$$(4.10) \qquad c_0 = \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} (T_i \bar{T}_i^{-1} T_i^* w_i, w_i) < \infty \text{ with } w_i = \sum_{j=i}^{J} v_j - T_i^{-1} v_i.$$

Proof. The proof of the theorem is based on a sequence of identities. We first set

$$(4.11) \qquad E_0 = I \text{ and } E_i = (I - T_i)E_{i-1} \text{ for } i = 1 : J.$$

We then have, with $E_J = E$,

$$
\begin{aligned}
\|v\|^2 - \|Ev\|^2 &= \sum_{i=1}^{J} \left(\|E_{i-1}v\|^2 - \|E_i v\|^2 \right) \\
&= \sum_{i=1}^{J} \left((E_{i-1}v, E_{i-1}v) - ((I - T_i)E_{i-1}v, (I - T_i)E_{i-1}v) \right) \\
&= \sum_{i=1}^{J} ((I - (I - T_i)^*(I - T_i))E_{i-1}v, \ E_{i-1}v) \\
&= \sum_{i=1}^{J} (\bar{T}_i E_{i-1}v, E_{i-1}v).
\end{aligned}
$$

Namely

$$(4.12) \qquad \|v\|^2 - \|Ev\|^2 = \sum_{i=1}^{J} (\bar{T}_i E_{i-1}v, E_{i-1}v).$$

This identity is instrumental in many existing convergence analyses for the method of subspace corrections (see Bramble, Pasciak, Wang and Xu [9], Bramble [8], Wang [32] and Xu [35]), and it is the starting point of our analysis here.

We now consider the product space $V^J = V \times V \times \cdots \times V$ and $\widetilde{V} = V_1 \times V_2 \times \cdots \times V_J \subset V^J$. We write the elements in this product space as column vectors:

$$
\tilde{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_J \end{pmatrix}, \quad \tilde{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_J \end{pmatrix}, \qquad u_i, v_i \in V \ (i = 1 : J), \quad \tilde{u}, \tilde{v} \in V^J,
$$

and use the inner product in the usual way:

$$
(\tilde{u}, \tilde{v})_{V^J} = \sum_{i=1}^{J} (u_i, v_i)_H.
$$

We introduce the following operators:

$$
\underset{\sim}{I} = \begin{pmatrix} I \\ I \\ \vdots \\ I \end{pmatrix}, \quad \underset{\sim}{E} = \begin{pmatrix} I \\ E_1 \\ \vdots \\ E_{J-1} \end{pmatrix}, \quad \underset{\sim}{L} = \begin{pmatrix} I & 0 & 0 & \ldots & 0 \\ T_1 & I & 0 & \ldots & 0 \\ T_1 & T_2 & I & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_1 & T_2 & T_3 & \ldots & I \end{pmatrix}
$$

and $\underset{\sim}{\bar{T}} = \mathrm{diag}(\bar{T}_1, \bar{T}_2, \ldots, \bar{T}_J)$.

Note that $\underset{\sim}{I}, \underset{\sim}{E} : V \mapsto V^J$, $\underset{\sim}{L} : V^J \mapsto V^J$ and $\underset{\sim}{\bar{T}} : V^J \mapsto \widetilde{V}$. Furthermore $\underset{\sim}{L} : V^J \mapsto V^J$ and $\underset{\sim}{\bar{T}} : \widetilde{V} \mapsto \widetilde{V}$ are apparently isomorphisms.

By the defining relation (4.11) for E_i, we have the identities

$$\sum_{j=1}^{i-1} T_j E_{j-1} + E_{i-1} = I, \quad i = 1, 2, \ldots, J,$$

which can be written in the following compact form:

$$\underset{\sim}{\underline{L}}\,\underset{\sim}{\underline{E}} = \underset{\sim}{\underline{I}}.$$

In view of (4.12), we have

$$
\begin{aligned}
\|v\|^2 - \|Ev\|^2 &= (\bar{\underset{\approx}{\underline{T}}}\,\underset{\sim}{\underline{E}}v, \underset{\sim}{\underline{E}}v)_{V^J} \\
&= (\bar{\underset{\approx}{\underline{T}}}\,\underset{\sim}{\underline{L}}^{-1}\underset{\sim}{\underline{I}}v, \underset{\sim}{\underline{L}}^{-1}\underset{\sim}{\underline{I}}v)_{V^J} \\
&= (\underset{\sim}{\underline{I}}^*(\underset{\sim}{\underline{L}}^*)^{-1}\bar{\underset{\approx}{\underline{T}}}\,\underset{\sim}{\underline{L}}^{-1}\underset{\sim}{\underline{I}}v, v).
\end{aligned}
$$

Namely

$$(4.13) \qquad \|v\|^2 - \|Ev\|^2 = (\underset{\sim}{\underline{I}}^*(\underset{\sim}{\underline{L}}^*)^{-1}\bar{\underset{\approx}{\underline{T}}}\,\underset{\sim}{\underline{L}}^{-1}\underset{\sim}{\underline{I}}v, v).$$

The derivation so far has been rather straightforward by using properly chosen notation.

We proceed to further modify (4.13). Let $\underset{\approx}{\underline{T}} = \mathrm{diag}(T_1, T_2, \ldots, T_J) : \widetilde{V} \mapsto \widetilde{V}$. We note that

$$(4.14) \qquad [(\underset{\sim}{\underline{L}}^*\underset{\approx}{\underline{T}} - \bar{\underset{\approx}{\underline{T}}})\tilde{v}]_i = T_i^*(\sum_{j=i}^{J} T_j v_j - v_i)$$

which, thanks to (3) of Lemma 4.1, implies that $\mathcal{R}(\underset{\sim}{\underline{L}}^*\underset{\approx}{\underline{T}} - \bar{\underset{\approx}{\underline{T}}}) \subset \widetilde{V}$. Since $\bar{\underset{\approx}{\underline{T}}} : \widetilde{V} \mapsto \widetilde{V}$ is an isomorphism, we can define the following operator (from $\widetilde{V}$ to $\widetilde{V}$):

$$\underset{\approx}{\underline{S}} = (\underset{\approx}{\underline{T}}^*\underset{\sim}{\underline{L}} - \bar{\underset{\approx}{\underline{T}}})\bar{\underset{\approx}{\underline{T}}}^{-1}(\underset{\sim}{\underline{L}}^*\underset{\approx}{\underline{T}} - \bar{\underset{\approx}{\underline{T}}}).$$

By (4.14), we have

$$(4.15) \qquad (\underset{\approx}{\underline{S}}\tilde{v}, \tilde{v}) = \sum_{i=1}^{J}(\bar{T}_i^{-1}T_i^*u_i, T_i^*u_i) \text{ with } u_i = \sum_{j=i}^{J} T_j v_j - v_i.$$

By means of the operator $\underset{\approx}{\underline{S}}$, we claim that the following relation holds:

$$(4.16) \qquad \|v\|^2 - \|Ev\|^2 = \left(\underset{\sim}{\underline{T}}(\underset{\approx}{\underline{S}} + \underset{\sim}{\underline{T}}^*\underset{\sim}{\underline{T}})^{-1}\underset{\sim}{\underline{T}}^*v, v \right)$$

where, with a slight abuse of notation,

$$(4.17) \qquad \underset{\sim}{\underline{T}} = (\underset{\sim}{\underline{T}}^*)^* = (T_1, \ldots, T_J) : \widetilde{V} \mapsto V.$$

The identity (4.16) is one crucial step in our derivation. Its discovery might not have been so obvious, but its verification can be carried out by direct calculations.

First, we have

$$\begin{aligned}
\underset{\sim}{S} &= (\underset{\sim}{T}^*\underset{\sim}{L}\bar{\underset{\sim}{T}}^{-1} - \underset{\sim}{I})(\underset{\sim}{L}^*\underset{\sim}{T} - \bar{\underset{\sim}{T}})\\[4pt]
&= \underset{\sim}{T}^*\underset{\sim}{L}\bar{\underset{\sim}{T}}^{-1}\underset{\sim}{L}^*\underset{\sim}{T} - \underset{\sim}{T}^*\underset{\sim}{L} - \underset{\sim}{L}^*\underset{\sim}{T} + \bar{\underset{\sim}{T}}\\[4pt]
&= \underset{\sim}{T}^*\underset{\sim}{L}\bar{\underset{\sim}{T}}^{-1}\underset{\sim}{L}^*\underset{\sim}{T} - \underset{\sim}{T}^*\underset{\sim}{T},
\end{aligned}$$

since $\underset{\sim}{T}^*\underset{\sim}{T} = \underset{\sim}{T}^*\underset{\sim}{L} + \underset{\sim}{L}^*\underset{\sim}{T} - \bar{\underset{\sim}{T}}$. Thus

$$\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T} = \underset{\sim}{T}^*\underset{\sim}{L}\bar{\underset{\sim}{T}}^{-1}\underset{\sim}{L}^*\underset{\sim}{T}$$

and hence

$$\begin{aligned}
\underset{\sim}{I}^*(\underset{\sim}{L}^*)^{-1}\bar{\underset{\sim}{T}}\underset{\sim}{L}^{-1}\underset{\sim}{I} &= \underset{\sim}{I}^*(\underset{\sim}{L}\bar{\underset{\sim}{T}}^{-1}\underset{\sim}{L}^*)^{-1}\underset{\sim}{I}\\[4pt]
&= \underset{\sim}{I}^*\underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*\underset{\sim}{I}\\[4pt]
&= \underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*.
\end{aligned}$$

This leads to the desired identity (4.16).

By the identity (4.16), we deduce that

$$(4.18)\qquad \|E\|^2 = 1 - \inf_{\|v\|=1}\left(\underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*v, v\right).$$

In regard to the last term in the above identity, we claim that the following identity holds:

$$(4.19)\qquad \inf_{\|v\|=1}\left(\underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*v, v\right) = \frac{1}{1 + c_0}$$

where

$$(4.20)\qquad c_0 = \sup_{\|v\|=1}\ \inf_{\underset{\sim}{T}\tilde{v}=v}(\underset{\sim}{S}\tilde{v}, \tilde{v}).$$

Let us now prove (4.19). To proceed, for any $w \in V$, let

$$\tilde{w} = (\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w, \quad v = \underset{\sim}{T}\tilde{w}.$$

By a simple calculation, we have

$$\frac{(\underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w, w)}{(\underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w, \underset{\sim}{T}(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w)} = \frac{(\underset{\sim}{S}\tilde{w}, \tilde{w}) + \|v\|^2}{\|v\|^2}.$$

By writing $\underset{\sim}{S} = \underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T} - \underset{\sim}{T}^*\underset{\sim}{T}$, it is easy to verify that

$$(\underset{\sim}{S}\tilde{w}, \tilde{\phi}) = 0, \quad \forall\tilde{\phi} \in \mathcal{N}(\underset{\sim}{T})$$

which implies that

$$(\underset{\sim}{S}\tilde{w}, \tilde{w}) = \inf_{\underset{\sim}{T}\tilde{v}=v}(\underset{\sim}{S}\tilde{v}, \tilde{v}).$$

Using the fact that $\underset{\sim}{T}$ is onto and therefore $\underset{\sim}{T}^*$ is one-to-one, we conclude that $\underset{\sim}{T}(\underset{\approx}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*$ is a symmetric, positive definite isomorphism. Hence

$$\left(\inf_{\|v\|=1}(\underset{\sim}{T}(\underset{\approx}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*v, v)\right)^{-1}$$
$$= \sup_{w\in V}\frac{(\underset{\sim}{T}(\underset{\approx}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w, w)}{(\underset{\sim}{T}(\underset{\approx}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w, \underset{\sim}{T}(\underset{\approx}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*w)}$$
$$= \sup_{w\in V}\inf_{\underset{\sim}{T}\tilde{v}=v}\frac{(\underset{\approx}{S}\tilde{v}, \tilde{v})}{\|v\|^2} + 1 = \sup_{v\in V}\inf_{\underset{\sim}{T}\tilde{v}=v}\frac{(\underset{\approx}{S}\tilde{v}, \tilde{v})}{\|v\|^2} + 1 = c_0 + 1.$$

This proves the identity (4.19).

At this point, we can conclude that the identity (4.9) can be obtained by combining (4.18), (4.19) and (4.15) together with a simple change of variable $T_i v_i \leftrightarrow v_i$.

Finally, thanks to (2.15), we can easily prove that $c_0 < \infty$ (see §4.5 for relevant details). This completes the proof. $\qquad\square$

Remark. In some sense, our (new) identity (4.9) for the method of successive subspace corrections is analogous to the (known) simple identity (2.17) for the method of parallel subspace corrections although (4.9) is much more nontrivial than (2.17).

As a special case of Theorem 4.2, we have the following result when all the subspace solvers are exact.

Corollary 4.3. *Under the assumption* (**A0**), *the following identity holds:*

$$(4.21) \qquad \|E\|^2 \equiv \|(I - P_J)(I - P_{J-1})\cdots(I - P_1)\|^2 = \frac{c_0}{1 + c_0}$$

where

$$(4.22) \qquad c_0 = \sup_{\|v\|=1}\inf_{\underset{\sim}{T}\tilde{v}=v}\sum_{i=1}^{J}\|P_i\sum_{j=i+1}^{J}v_j\|^2.$$

4.3. **Application to the method of alternating projections.** In this section, we shall study the product operator E as an operator on the space H which may contain V as a proper subspace. This line of result is related to the convergence of the method of alternating projections.

Introducing the notation

$$\mathrm{Fix}(E) = \mathcal{N}(I - E) = \{v \in H \ : \ Ev = v\}$$

we have the following simple result on $\mathrm{Fix}(E)$:

Lemma 4.4. *Let* $E = (I - T_J)\cdots(I - T_1)$. *If each* T_i *satisfies* (**A1**) *and* (**A2**), *then*

$$\mathrm{Fix}(E) = \mathrm{Fix}(E^*) = \bigcap_{i=1}^{J}\mathrm{Fix}(I - T_i) = \bigcap_{i=1}^{J}\mathcal{N}(T_i) = \bigcap_{i=1}^{J}V_i^{\perp}$$

and, with $M = \mathrm{Fix}(E)$,

$$V \equiv M^{\perp} = \overline{\sum_{i=1}^{J}(\mathcal{N}(T_i))^{\perp}} = \overline{\sum_{i=1}^{J}V_i}.$$

Proof. We first prove that

$$(4.23) \qquad \text{Fix}(E) = \bigcap_{i=1}^{J} \mathcal{N}(T_i).$$

If $v \in \mathcal{N}(T_i)$, for all $i = 1 : J$, then obviously $Ev = v$ hence $v \in \text{Fix}(E)$. This proves that $\text{Fix}(E) \supset \bigcap_{i=1}^{J} \mathcal{N}(T_i)$. To prove the other inclusion, let $v \in \text{Fix}(E)$. Since every $I - T_i$ is nonexpansive, we have that

$$\|v\|^2 = \|Ev\|^2 \le \|(I - T_1)v\|^2.$$

By Lemma 4.1 we obtain that $v \in \mathcal{N}(T_1)$. A simple induction argument shows that $v \in \mathcal{N}(T_i)$, for all $i = 1 : J$, that is

$$\text{Fix}(E) \subset \bigcap_{i=1}^{J} \mathcal{N}(T_i)$$

and (4.23) follows. The proof of the lemma can be easily completed by using (4.23) in combination with Lemma 4.1 and Lemma 3.2. $\qquad\blacksquare$

Combining Theorem 4.2 with the obvious identities

$$(4.24) \qquad E^n - P_M = (E - P_M)^n = (EP_V), \quad \|EP_V\| = \|E\|_{\mathcal{L}(V,V)}$$

gives the following theorem for the method of alternating projections.

Theorem 4.5. *Under the assumptions* (**A1**), (**A2**) *and* (**A0**),

$$(4.25) \qquad \|E(I - P_M)\|^2 = \|E - P_M\|^2 = \|EP_V\|^2 = \frac{c_0}{1 + c_0}$$

where c_0 is given by (4.10), *namely*

$$c_0 = \sup_{\|v\|=1} \inf_{\sum_{i=1}^{J} v_i = v} \sum_{i=1}^{J} (T_i \bar{T}_i^{-1} T_i^* w_i, w_i) \ \text{with} \ w_i = \sum_{j=i}^{J} v_j - T_i^{-1} v_i.$$

Consequently

$$(4.26) \qquad \|E^n v - P_M v\| \le (1 + c_0^{-1})^{-\frac{n}{2}} \|v\| \quad \forall v \in H.$$

Remark. We note that the above theorem implies the convergence of a generalization of the method of alternating projections, namely the projection on each subspace V_i only needs to be computed approximately: $T_i \approx P_{V_i}$ satisfying (**A2**).

4.4. **Qualitative convergence without assumption** (**A0**)**.** In this section, we shall present some results without assuming (**A0**). In this case, we can only obtain a qualitative convergence result as the following simple result indicates that (**A0**) is necessary for any quantitative result as in Theorem 4.2.

Theorem 4.6. *Assume that* (**A1**) *and* (**A2**) *hold. Then the following two statements are equivalent:*

(1) $\sum_{i=1}^{J} V_i$ *is closed, namely* (**A0**) *is satisfied.*
(2) $\|E\|_{\mathcal{L}(V,V)} < 1$.

Proof. By our main Theorem 4.2, (1) implies (2). Now, if (2) holds, then by (4.16) and the obvious fact that $\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T}$ is isomorphism on $\tilde{V}$,

$$
\begin{aligned}
(1 - \delta^2)\|v\|^2 &\leq \|v\|^2 - \|Ev\|^2 \\
&= \left((\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\underset{\sim}{T}^*v, \underset{\sim}{T}^*v \right) \\
&\leq \|(\underset{\sim}{S} + \underset{\sim}{T}^*\underset{\sim}{T})^{-1}\|(\underset{\sim}{T}\underset{\sim}{T}^*v, v).
\end{aligned}
$$

This means that $\underset{\sim}{T}\underset{\sim}{T}^* = \sum_{i=1}^J T_i T_i^*$ is a symmetric positive definite isomorphism on V which, by Theorem 2.3, implies that $V = \sum_i V_i$, namely (1) holds. $\qquad\square$

We would like to point out that the above result was known before in the case that each T_i is a projection; see Bauschke, Borwein and Lewis [3].

In view of Theorem 2.3 and Theorem 4.6, the convergence behavior of the method of successive subspace corrections should be related to that of the method of parallel subspace corrections. In fact, the norm $\|E\|$ and the extreme eigenvalues of T may be estimated in terms of each other and some estimates of this type can be deduced from the proof of Theorems 2.3 and 4.6. We also refer to Griebel and Oswald [23] and Bramble and Zhang [13] for an estimate of $\|E\|$ in terms of the condition number of T.

We are now in a position to present the main result of this section.

Theorem 4.7. *Let $E = (I - T_J) \cdots (I - T_1)$. If each T_i satisfies* (**A1**) *and* (**A2**), *then*

$$(4.27) \qquad \lim_{n \to \infty} \|E^n v - P_M v\| = 0, \quad \forall v \in H,$$

or

$$\lim_{n \to \infty} \|[E(I - P_M)]^n v\| = 0, \quad \forall v \in H.$$

Proof.

$$
\begin{aligned}
\|v - Ev\|^2 &= \|\sum_{i=1}^J (E_{i-1}v - E_i v)\|^2 \\
&\leq J \sum_{i=1}^J \|E_{i-1}v - E_i v\|^2 \quad \text{(Cauchy-Schwarz inequality)} \\
&= J \sum_{i=1}^J \|T_i E_{i-1}v\|^2 \\
&\leq \frac{J\omega}{2 - \omega} \sum_{i=1}^J (\|E_{i-1}v\|^2 - \|E_i v\|^2) \quad \text{((4) of Lemma 4.1)} \\
&= \frac{J\omega}{2 - \omega}(\|v\|^2 - \|Ev\|^2).
\end{aligned}
$$

Replacing v by $E_J^n v$, we have

$$\|E^n(I - E)v\|^2 \leq \frac{J\omega}{2 - \omega}(\|E^n v\|^2 - \|E^{n+1}v\|^2) \to 0$$

as $n \to \infty$ because $\|E^n v\|^2$, as a nonnegative decreasing sequence (since $\|E\| \leq 1$), has a finite limit. Thus $E^n v \to 0$ for any $v \in \mathcal{R}(I - E)$. But since $\|E\| \leq 1$, we conclude that $E^n v \to 0$ for all

$$v \in \overline{\mathcal{R}(I - E)} = \mathcal{N}(I - E^*)^\perp = (\text{Fix}(E^*))^\perp = M^\perp.$$

Thus, for any $v \in H$, we have

$$E^n v - P_M v = E^n (I - P_M) v = E^n (P_M^\perp v) \to 0$$

as desired. □

The above theorem generalizes a result, due to Smarzewski [28] (see Bauschke, Deutsch, Hundal and Park [4]), which states that (4.27) holds if each T_i is *self-adjoint, nonnegative* and *nonexpansive* (namely T_i is self-adjoint and satisfies (**A2**) with $\omega = 1$).

The above theorem is also related to a general result stated below (see Bauschke, Deutsch, Hundal and Park [4]).

Proposition 4.8. *Let E be a nonexpansive linear operator on H and $M \subset H$ a closed subspace. The following statements are equivalent:*

(1) $\lim_n \|E^n v - P_M v\| = 0$ *for all $v \in H$.*
(2) $M = \text{Fix}(E)$ *and $E^n v \to 0$ for all $v \in M^\perp$.*
(3) $M = \text{Fix}(E)$ *and E is "asymptotically regular", namely $E^n v - E^{n+1} v \to 0$ for all $v \in H$.*

4.5. **On the estimate of c_0.** We shall now briefly discuss how the constant c_0 in (4.9) can be estimated. Basically, most existing theories can be recovered and/or improved by certain upper-bound estimates of c_0, but we shall not go through all those estimates. Here we shall give some simple examples as an illustration. We shall consider the special case when $a(\cdot, \cdot)$ is a symmetric, positive, definite bilinear form, namely it is an inner product on H.

The following intriguing identity is useful.

Lemma 4.9. *The following identity holds, for all $i = 1 : J$:*

$$(4.28) \qquad (T_i^{*-1} - I) T_i \bar{T}_i^{-1} T_i^* (T_i^{-1} - I) = \bar{T}_i^{-1} - I.$$

Proof. Let $S_i = T_i^{*-1} + T_i^{-1} - I$. It follows that

$$
\begin{aligned}
&(T_i^{*-1} - I) T_i \bar{T}^{-1} T_i^* (T_i^{-1} - I) \\
&= (T_i^{*-1} - I)(T_i^{*-1} + T_i^{-1} - I)^{-1} (T_i^{-1} - I) \\
&= (S_i - T_i^{-1}) S_i^{-1} (S_i - (T_i^*)^{-1}) \\
&= S_i - T_i^{-1} - (T_i^*)^{-1} + T_i^{-1} S_i^{-1} (T_i^*)^{-1} \\
&= -I + \bar{T}_i^{-1}.
\end{aligned}
$$

□

A special decomposition. In some important applications, such as multigrid methods, a special decomposition $v = \sum_i v_i$ may be obtained by a sequence of linear operators. If we assume that there exist operators: $\Pi_i : V \mapsto V_i$ such that

$$\mathcal{R}(\Pi_i - \Pi_{i-1}) \subset V_i, \quad \forall v \in V,$$

with $\Pi_0 = 0$ and $\Pi_J = I$, we have a telescopic decomposition

$$v = \sum_{i=1}^{J} v_i \quad \text{with} \quad v_i = (\Pi_i - \Pi_{i-1})v.$$

Then, in view of (4.10), we have

$$c_0 \le \sup_{\|v\|=1} \sum_{i=1}^{J} (T_i \bar{T}_i^{-1} T_i^* w_i, w_i)$$

with

(4.29) $$w_i = \sum_{j=i}^{J} v_j - T_i^{-1} v_i = (I - \Pi_i)v + (I - T_i^{-1})(\Pi_i - \Pi_{i-1})v.$$

Furthermore

$$\begin{aligned} c_0 \quad &\le \quad \sup_{\|v\|=1} \sum_{i=1}^{J} (T_i \bar{T}_i^{-1} T_i^* w_i, w_i) \\ &\le \quad 2 \sup_{\|v\|=1} \sum_{i=1}^{J} \bigg((T_i \bar{T}_i^{-1} T_i^*(I - \Pi_i)v, (I - \Pi_i)v) \\ & \qquad\qquad\qquad + ((\bar{T}_i^{-1} - I)(\Pi_i - \Pi_{i-1})v, (\Pi_i - \Pi_{i-1})v) \bigg). \end{aligned}$$

An important special case is when $\Pi_i = P_i$. In this case, the expression (4.29) is reduced to

$$w_i = \sum_{j=i}^{J} (I - T_i^{-1})(P_i - P_{i-1})v$$

and

(4.30) $$c_0 \le \sup_{\|v\|=1} \sum_{i=1}^{J} ((\bar{T}_i^{-1} - I)(P_i - P_{i-1})v, (P_i - P_{i-1})v).$$

A general estimate. We now turn to another extreme to obtain some conservative estimates in the most general setting.

Let $\bar{\underset{\sim}{G}} = \text{diag}(\bar{T}_1, \dots, \bar{T}_J)$ and

$$\underset{\sim}{T} = \begin{pmatrix} T_1^*(T_1^{-1} - I)P_1 & T_1^* & T_1^* & \cdots & T_1^* \\ 0 & T_2^*(T_2^{-1} - I)P_2 & T_2^* & \cdots & T_2^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & T_J^*(T_J^{-1} - I)P_J \end{pmatrix}.$$

Then

$$\begin{aligned} c_0 \quad &= \quad \sup_{\|v\|=1} \inf_{\sum \bar{T}_i v_i = v} \|\bar{\underset{\sim}{G}}^{-\frac{1}{2}} \underset{\sim}{T} \bar{\underset{\sim}{G}}^{\frac{1}{2}} (\bar{\underset{\sim}{G}}^{\frac{1}{2}} \tilde{v})\|^2 \\ &\le \quad \sup_{\|v\|=1} \inf_{\sum \bar{T}_i v_i = v} \|\bar{\underset{\sim}{G}}^{-\frac{1}{2}} \underset{\sim}{T} \bar{\underset{\sim}{G}}^{\frac{1}{2}}\|^2 \|\bar{\underset{\sim}{G}}^{\frac{1}{2}} \tilde{v}\|^2 \\ &= \quad K_0 K_1 \end{aligned}$$

where

$$K_0 = \sup_{\|v\|=1} \inf_{\sum \bar{T}_i v_i = v} \|\bar{G}^{\frac{1}{2}} \tilde{v}\|^2 = \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{k=1}^{J} (\bar{T}_k^{-1} v_k, v_k)$$

and

$$K_1 = \|\underset{\sim}{\bar{G}}^{-\frac{1}{2}} \underset{\sim}{T} \underset{\sim}{\bar{G}}^{\frac{1}{2}}\|^2.$$

For the special case that $T_i = P_i$, namely all subspace solvers are exact, we have

$$c_0 = \sup_{\|v\|=1} \inf_{\sum v_i = v} \sum_{k=1}^{J} \|P_k \sum_{i=k+1}^{J} v_i\|$$

and the matrix $\underset{\sim}{T}$ is reduced to

$$\underset{\sim}{P} = \begin{pmatrix} 0 & P_1 & P_1 & \dots & P_1 & P_1 \\ 0 & 0 & P_2 & \dots & P_2 & P_2 \\ 0 & 0 & 0 & \dots & P_3 & P_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & P_{J-1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus

$$c_0 \le K_0 K_1$$

with

$$K_0 = \sup_{\|v\|=1} \inf_{\sum_i v_i = v} \sum_{i=1}^{J} \|v_i\|^2, \quad K_1 = \|\underset{\sim}{P}\|^2.$$

5. Some applications

In this last section, we shall give some examples to demonstrate how our main result Theorem 4.2 can be applied in concrete situations. We shall consider two simple examples, one is an overlapping domain decomposition method (with exact subspace solvers) and the other is a multigrid method (with inexact subspace solvers). The relevant convergence results are mostly known in the literature but they will be obtained here in a different fashion by using our new theory.

We consider the following simple problem: Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = f(v), \quad \forall v \in H_0^1(\Omega).$$

Here $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain and

$$a(u,v) = \int_\Omega \left(\sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \right) dx, \quad f(v) \equiv \int_\Omega fv \, dx.$$

5.1. **An overlapping domain decomposition method.** We shall now describe a simple overlapping domain decomposition method for solving (5.4) and analyze its convergence property using our new theory. We shall only make a very weak assumption on the coefficients (a_{ij}): namely we assume that each a_{ij} is bounded and measurable (namely $a_{ij} \in L^\infty(\Omega)$) and there exist two positive constants Λ_0 and Λ_1 such that

$$(5.1) \qquad \Lambda_0 \sum_{i=1}^d |\xi_i|^2 \le \sum_{i,j=1}^d a_{ij}(x)\xi_i\xi_j \le \Lambda_1 \sum_{i=1}^d |\xi_i|^2, \quad \forall \xi = (\xi_i) \in R^d, x \in \Omega.$$

We assume that we are given a set of overlapping subdomains $\{\Omega_i\}_{i=1}^J$ of Ω such that $\Omega = \sum_{i=1}^J \Omega_i$. While there are many possible ways of defining such a decomposition, as an illustration let us now give a specific example. We start by assuming that Ω is triangulated with $\Omega = \bigcup_{i=1}^J \tau_i$, where the τ_i's are nonoverlapping simplexes of size h_0, with $h_0 \in (0, 1]$ and quasi-uniform, i.e. there exist constants C_0 and C_1 not depending on h such that each simplex τ_i is contained in (contains) a ball of radius $C_1 h_0$ (respectively $C_0 h_0$). Given every grid point x_i in this triangulation, we can define a subdomain Ω_i to be the union of all simplexes containing x_i as a vertex. We then obtain an overlapping decomposition $\Omega = \sum_{i=1}^J \Omega_i$.

Let $V = H_0^1(\Omega)$. Associated with each subdomain, we introduce the subspaces $V_i \subset V$ ($1 \le i \le J$) as follows:

$$V_i = \{v \in V : v(x) = 0, \ \forall x \in \Omega \setminus \Omega_i\}.$$

We further introduce a linear finite element subspace V_0 defined from the aforementioned quasi-uniform triangulation of Ω, namely

$$V_0 = \{v \in H_0^1(\Omega) : v|_\tau \in \mathcal{P}_1(\tau_i), \ \forall \tau_i\}$$

where $\mathcal{P}_1$ is the space of linear polynomials.

With this definition of subspaces, a successive subspace correction method can naturally be defined. This type of algorithm is often known as the Schwarz overlapping domain decomposition method. In this example, we consider the case that all subspace solvers are exact. But of course, the case of inexact subspace solvers can also be treated without many more difficulties.

Let us now proceed with a concise convergence analysis for this overlapping Schwarz method (with exact subspace solvers). As in the existing literature, one main ingredient of the analysis is a partition of unity, $\{\theta_i\}_{i=1}^J$, defined on Ω satisfying $\sum_{i=1}^J \theta_i = 1$ and, for $i = 1, \dots, J$,

$$\mathrm{supp}\theta_i \subset \Omega_i \cup \partial\Omega, \quad 0 \le \theta_i \le 1, \quad \max_{x \in \overline{\Omega}_i} |\nabla\theta_i(x)| \le c_1 h_0^{-1}.$$

The construction of such a partition of unity is standard. In this specific example, we may simply choose each θ_i to be the local nodal basis function associated with the grid x_i. Another ingredient in our analysis is the L^2 projection $Q_0 : V \mapsto V_0$ which is known to satisfy (see Bramble and Xu [12])

$$h_0^{-1}\|v - Q_0 v\|_{0,\Omega} + |v - Q_0 v|_{1,\Omega} \le c_2 |v|_{1,\Omega}.$$

With the partition of unity (θ_i) and the L^2 projection Q_0, given any $v \in V$, a partition $v = \sum_{i=0}^J v_i$ for $v_i \in V_i$ can then be obtained with

$$(5.2) \qquad\qquad v_0 = Q_0 v, \quad v_i = \theta_i(v - Q_0 v), \quad i = 1, \dots, J,$$

where $Q_0 : V \mapsto V_0$ is the L^2 projection.

Introducing the notation

$$\|v\|_{A,\Omega}^2 = \int_\Omega \sum_{i,j=1}^d a_{ij}(x) \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j},$$

we have

$$\sum_{k=0}^J \|P_k \sum_{i=k+1}^J v_i\|_{A,\Omega}^2 \le \|v - Q_0 v\|_{A,\Omega}^2 + \sum_{k=1}^J \|(\sum_{i=k+1}^J \theta_i)(v - Q_0 v))\|_{A,\Omega_k}^2$$

$$\le \Lambda_1 \left(|v - Q_0 v|_{1,\Omega}^2 + \sum_{k=1}^J \max_{x \in \overline{\Omega}_k} | \sum_{i=k+1}^J \nabla \theta_i(x)| \|v - Q_0 v\|_{0,\Omega_k}^2 + |v - Q_0 v|_{1,\Omega_k}^2 \right)$$

$$\le \Lambda_1 C_0 |v|_{1,\Omega}^2 \le \frac{\Lambda_1}{\Lambda_0} C_0 \|v\|_{A,\Omega}^2$$

with C_0 dependent of c_1, c_2 (and the maximal number of subdomains to which a given point may belong), but independent of a_{ij}, h_0 and J. Hence for c_0 given by (4.22), we have $c_0 \le \frac{\Lambda_1}{\Lambda_0} C_0$.

We conclude from the above analysis that the classic overlapping Schwarz domain decomposition method converges uniformly for the elliptic boundary value problem with general bounded measurable coefficients and the rate of convergenc only depends on the bounds, Λ_0 and Λ_1, of the extreme eigenvalues of the coefficient matrix $(a_{ij}(x))$.

One interesting case is when the coefficients a_{ij} are highly oscillatory and, as we see, the convergence rate of the classic Schwarz overlapping domain decomposition is not affected by these possible oscillations as long as the a_{ij} stay within the bounds as in (5.1).

By using the new theory, our analysis presented above is straightforward and transparent. For works related to this problem, we refer to [19, 20, 35, 29].

5.2. **A multigrid method.** In this subsection we shall give another example of the application of our identity to obtain a convergence estimate for a multigrid algorithm. This estimate was known before, but the proof here is different from those in the literature (see Braess and Hackbusch [6] and Bramble and Pasciak [7]).

We assume that Ω has been triangulated with a nested sequence of quasi-uniform triangulations $T_k = \{\tau_k^i\}$ of size h for $k = 0, \dots, j$ where the quasi-uniformity constants are independent of k. These triangulations should be nested in the sense that any triangle τ_{k-1}^l can be written as a union of triangles of $\{\tau_k^i\}$. We further assume that there is a constant $\gamma > 1$, independent of k, such that

$$h_k \sim \gamma^{-k}.$$

Associated with each T_k, a finite element space $V_k \subset H_0^1(\Omega)$ can be defined. One has

$$(5.3) \qquad V_0 \subset V_1 \subset \cdots \subset V_k \subset \cdots \subset V_J.$$

We are interested in the multigrid method for solving the following finite element equation: Find $u_h \in V \equiv V_J$ satisfying

$$(5.4) \qquad a(u_h, v) = f(v), \quad \forall v \in V,$$

where

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v.$$

Define $A_i : V_i \mapsto V_i$ by:

$$(A_i u_i, v_i)_0 = a(u_i, v_i), \quad \forall u_i, v_i \in V_i.$$

The subspace solver T_i in a multigrid method is often given by applying a number of, say m, smoothings using, for example, a local relaxation method such as Gauss-Seidel iteration (with perhaps the exception of $i = 0$ in which an exact solver may be used, namely $T_0 = P_0$). Let $R_i \approx A_i^{-1}$ denote one such smoothing. With m smoothings, one strategy is to apply R_i and R_i^t alternatively. Here R_i^t is the adjoint of R_i with respect to $(\cdot, \cdot)$; for example, if R_i represents forward Gauss-Seidel iteration, then R_i^t represents backward Gauss-Seidel iteration. With this kind of subspace correction, we have

$$T_i = \begin{cases} P_i - (K_i^* K_i)^{\frac{m}{2}} P_i, & \text{if } m \text{ is even,} \\ P_i - K_i (K_i^* K_i)^{\frac{m-1}{2}} P_i, & \text{if } m \text{ is odd,} \end{cases}$$

and

$$\bar{T}_i = (I - K_{i,m}^m) P_i$$

where

$$K_i = I - R_i A_i, \quad K_i^* = I - R_i^t A_i$$

and

$$K_{i,m} = \begin{cases} K_i^* K_i, & \text{if } m \text{ is even,} \\ K_i K_i^*, & \text{if } m \text{ is odd.} \end{cases}$$

Note that K_i^* is the adjoint of K_i with respect to $a(\cdot, \cdot)$.

In this example, we shall make two assumptions. The first assumption is on the approximation of the finite element subspaces:

$$(5.5) \qquad \|(I - P_{i-1}) v_i\|^2 \le \frac{c_1}{\lambda_i} a(v_i, v_i), \quad \forall v \in V_i,$$

where $\lambda_i = \rho(A_i)$. The second assumption is on the smoother R_i:

$$(5.6) \qquad \frac{c_2}{\lambda_i}(v, v) \le (\bar{R}_i v, v), \quad \forall v \in V_i,$$

where $\bar{R}_i$ is the corresponding symmetrization of R_i.

We note that the first assumption is satisfied, for example, when a_{ij} are smooth and Ω is smooth or a convex Lipschitz domain, and the second assumption is satisfied for Gauss-Seidel iteration or the damped Jacobi method (see Xu [35])

To estimate the constant c_0, we consider the decomposition $v = \sum_i v_i$ for any $v \in V$ with

$$(5.7) \qquad v_i = (P_i - P_{i-1}) v.$$

Since, thanks to (5.3), $P_{i-1} = P_{i-1} P_i = P_i P_{i-1}$, (5.5) implies that

$$(5.8) \qquad \lambda_i(v_i, v_i) \le c_1 a(v_i, v_i).$$

We note that

$$
\begin{aligned}
a(\bar{T}_i^{-1}(I - \bar{T}_i)v_i, v_i) &= a((I - K_{i,m}^m)^{-1}K_{i,m}^m v_i, v_i) \\
&= (\bar{R}_i^{-1}\bar{R}_i A_i (I - K_{i,m}^m)^{-1}K_{i,m}^m v_i, v_i) \\
&\leq \frac{\lambda_i}{c_2}((I - K_{i,m})(I - K_{i,m}^m)^{-1}K_{i,m}^m v_i, v_i) \\
&\leq \frac{\lambda_i}{c_2}\max_{t \in [0,1]}[(1 - t^m)^{-1}t^m(1 - t)](v_i, v_i) \\
&\leq \frac{\lambda_i}{c_2 m}(v_i, v_i).
\end{aligned}
$$

Hence, by (5.8), we have

$$
\sum_{i=1}^{J} a(\bar{T}_i^{-1}(I - \bar{T}_i)v_i, v_i) \leq \sum_{i=1}^{J} \frac{\lambda_i}{c_2 m}(v_i, v_i) \leq \sum_{i=1}^{J} \frac{c_1}{c_2 m}a(v_i, v_i) = \frac{c_1}{c_2 m}a(v, v).
$$

Thus, by (4.30), we have

$$
c_0 \leq \frac{c_1}{c_2 m}.
$$

Consequently, the method of successive subspace corrections, based on multilevel subspaces (5.3) with a smoother satisfying (5.6), has the following convergence estimate:

$$
\frac{c_1}{c_1 + c_2 m}.
$$

Acknowledgment

References

1. N. Aronszjan, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404. MR **14:**479c

2. I. Babuška and A. K. Aziz, *Lectures on the mathematical foundations of the finite element method*, University of Maryland, College Park, Washington DC, 1972, Technical Note BN-748. MR **54:**9111

3. H. Bauschke, J. Borwein, and A. Lewis, *The method of cyclic projections for closed convex sets in Hilbert space*, Recent developments in optimization theory and nonlinear analysis (Jerusalem, 1995), Amer. Math. Soc., Providence, RI, 1997, pp. 1–38. MR **98c:**49069

4. H. Bauschke, F. Deutsch, H. Hundal, and S. Park, *Accelerating the convergence of the method of alternating projections*, Preprint, July 1999.

5. H. H. Bauschke and J. M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev. **38** (1996), no. 3, 367–426. MR **98f:**90045

6. D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V-cycle*, SIAM J. Numer. Anal. **20** (1983), no. 5, 967–975. MR **85h:**65233

7. J. Bramble and J. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp. **49** (1987), no. 180, 311–329. MR **89b:**65234

8. J. H. Bramble, *Multigrid methods*, Longman Scientific & Technical, Harlow, 1993. MR **95b:**65002

9. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. **57** (1991), no. 195, 23–45. MR **91m:**65158

10. _____, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp. **57** (1991), no. 195, 1–21. MR **92d:**65094

11. J. H. Bramble, J. E. Pasciak, and J. Xu, *Parallel multilevel preconditioners*, Math. Comp. **55** (1990), 1–22. MR **90k:**65170

12. J. H. Bramble and J. Xu, *Some estimates for a weighted L^2 projection*, Math. Comp. **56** (1991), no. 194, 463–476. MR **91k:**65140

13. J. H. Bramble and X. Zhang, *The analysis of multigrid methods*, Handbook of numerical analysis, Vol. VII, North-Holland, Amsterdam, 2000, pp. 173–415. MR **2001m:**65183

14. 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. MR **51:**1540

15. F. Deutsch, *Applications of von Neumann's alternating projections algorithm*, Mathematical methods in operations research (Sofia, 1983), Bulgar. Acad. Sci., Sofia, 1983, pp. 44–51. MR **86i:**90093

16. _____, *von Neumann's alternating method: the rate of convergence*, Approximation theory, IV (College Station, TX, 1983), Academic Press, New York, 1983, pp. 427–434. MR **85m:**41040

17. _____, *Rate of convergence of the method of alternating projections*, Parametric optimization and approximation (Oberwolfach, 1983), Birkhäuser, Basel, 1985, pp. 96–107. MR **88d:**41026

18. F. Deutsch, *The method of alternating orthogonal projections*, Approximation theory, spline functions and applications, Kluwer Acad. Publ., 1992, pp. 105–121. MR **93a:**41047

19. M. Dryja and O. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989), SIAM, Philadelphia, PA, 1990, pp. 3–21. MR **91m:**65294

20. _____, *Multilevel additive methods for elliptic finite element problems*, Parallel algorithms for partial differential equations (Kiel, 1990), Vieweg, Braunschweig, 1991, pp. 58–69. MR **93b:**65188

21. J. Gilbert and W. A. Light, *Multigrid methods and the alternating algorithm*, Algorithms for approximation (Shrivenham, 1985), Oxford Univ. Press, New York, 1987, pp. 447–458. MR **88j:**65116

22. M. Griebel and P. Oswald, *On additive Schwarz preconditioners for sparse grid discretizations*, Numer. Math. **66** (1994), no. 4, 449–463. MR **95a:**65188

23. _____, *On the abstract theory of additive and multiplicative Schwarz algorithms*, Numer. Math. **70** (1995), no. 2, 163–180. MR **96a:**65164

24. W. Hackbusch, *Multigrid methods and applications*, Springer-Verlag, Berlin, 1985. MR **87e:**65082

25. _____, *Iterative solution of large sparse systems of equations*, Springer-Verlag, New York, 1994. MR **94k:**65002

26. I. Halperin, *The product of projection operators*, Acta Sci. Math. (Szeged) **23** (1962), 96–99. MR **25:**5373

27. S. Kayalar and H. L. Weinert, *Error bounds for the method of alternating projections*, Math. Control Signals Systems **1** (1988), no. 1, 43–59. MR **89b:**65137

28. R. Smarzewski, *Iterative recovering of orthogonal projections*, Preprint, December 1996.

29. B. Smith, P. Bjørstad, and W. Gropp, *Domain decomposition*, Cambridge University Press, Cambridge, 1996, Parallel multilevel methods for elliptic partial differential equations. MR **98g:**65003

30. U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press Inc., San Diego, CA, 2001, With contributions by A. Brandt, P. Oswald and K. Stüben. MR **2002b:**65002

31. J. von Neumann, *The geometry of orthogonal spaces*, Functional operators- vol. II, Annals of Math. Studies, no. 22, Princeton University Press, 1950, This is a reprint of mimeographed lecture notes, first distributed in 1933. MR **11:**599e

32. J. Wang, *Convergence analysis without regularity assumptions for multigrid algorithms based on SOR smoothing*, SIAM J. Numer. Anal. **29** (1992), no. 4, 987–1001. MR **93e:**65141

33. O. B. Widlund, *Some Schwarz methods for symmetric and nonsymmetric elliptic problems*, Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations (Norfolk, VA, 1991), SIAM, Philadelphia, PA, 1992, pp. 19–36. MR **93j:**65202

34. J. Xu, *Theory of multilevel methods*, Ph.D. thesis, Cornell University, Ithaca, NY, 1989, AM report 48, Dept. of Math., Penn. State Univ., University Park, PA.

35. _____, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), 581–613. MR **93k:**65029

36. _____, *An introduction to multilevel methods*, Wavelets, multilevel methods and elliptic PDEs (Leicester, 1996), Oxford Univ. Press, New York, 1997, pp. 213–302. MR **99d:**65329
37. J. Xu and L. Zikatanov, *Some observations on Babuška and Brezzi theories*, Num. Math., 2000 (to appear).
38. J. Xu and J. Zou, *Some nonoverlapping domain decomposition methods*, SIAM Rev. **40** (1998), no. 4, 857–914. MR **99m:**65241
39. H. Yserentant, *Old and new convergence proofs for multigrid methods*, Acta numerica, 1993, Cambridge Univ. Press, Cambridge, 1993, pp. 285–326. MR **94i:**65128

CENTER FOR COMPUTATIONAL MATHEMATICS AND APPLICATIONS, DEPARTMENT OF MATHEMATICS, THE PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PENNSYLVANIA 16802
E-mail address: xu@math.psu.edu

CENTER FOR COMPUTATIONAL MATHEMATICS AND APPLICATIONS, DEPARTMENT OF MATHEMATICS, THE PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PENNSYLVANIA 16802
E-mail address: ltz@math.psu.edu