GLOBAL AND UNIFORM CONVERGENCE OF SUBSPACE CORRECTION METHODS FOR SOME CONVEX OPTIMIZATION PROBLEMS

XUE-CHEUNG TAI AND JINCHAO XU

ABSTRACT. This paper gives some global and uniform convergence estimates for a class of subspace correction (based on space decomposition) iterative methods applied to some unconstrained convex optimization problems. Some multigrid and domain decomposition methods are also discussed as special examples for solving some nonlinear elliptic boundary value problems.

1. INTRODUCTION

This paper is devoted to convergence analysis for a class of iterative methods for solving some convex optimization problems. It is well known that some iterative methods, such as Newton’s method, can be proven to be globally convergent to certain convex optimization problems. In this paper, we shall study the global convergence property of a class of iterative methods that include multigrid and domain decomposition methods.

Multigrid and domain decomposition methods have been studied extensively in recent years for linear partial differential equations. Recent research (see for example [37]) reveals that multigrid and domain decomposition methods can be described and analyzed under a general framework based on space decomposition and subspace correction (see also [3], [11], [27], [15], and [21]).

Naturally there is also a great deal of work on nonlinear problems. Some of these methods are more or less straightforward extensions from the ones for the linear problems, some of them are based on Newton’s method and the linearized problems are solved by linear methods. Rather than going into the details of various different techniques, let us just give a sample of references on these methods. For the work based on the linearization approach, we refer to Bank and Rose [2], Cai and Dryja [4], Rannacher [23], Deuhard and Weiser [8], Xu [38, 39], and Axelsson and Layton [1]. For the work based on multigrid or domain decomposition with nonlinear smoothers or nonlinear local solvers, we refer to Lions [19], Mandel [20], Gelman and Mandel [13], McCormick [21], Hackbusch and Reusken [16], Reusken [17].
Our algorithms bear some of the natures of the methods of Mandel [20], Gel-
man and Mandel [13], McCormick [21], Kornhuber [17, 18] in the sense that we are
reducing the original minimization problem into a number of smaller minimization
problems and trying to guarantee a monotone decreasing of the cost functional.
The nonlinear approach of Hackbusch and Reusken [16] and Reusken [24] diifers
from ours and the rate of convergence is in some sense local. The algorithm of Dryja
and Hackbusch [10] is the same as our parallel subspace correction algorithm, which
has also been studied earlier in [33, 30, 28], but our convergence results are quite
different. The convergence analysis presented here is valid for more general prob-
lems which can handle some nonlinear diffusion problems even when the nonlinear
diffusion coecient is degenerate or singular (see Section 5).

The iterative methods we will study in this paper can be viewed as a straight-
forward extension of the subspace correction iterative method for linear problems
as described in [37] in a similar manner as in [19, 28, 33]. Of course, in various
special applications (such as multigrid and domain decomposition methods), these
methods are either almost identical or very similar to some methods studied in
the aforementioned literature. The main concern of this paper is to establish some
global and uniform convergence estimates for a class of subspace correction iterative
methods for some unconstrained convex optimization problems. Some of the tech-
niques used in this paper are based on earlier works (28, see also 29, 30, 31,
32, 22 and 33). We would like to point out that most convergence estimates for
nonlinear problems in the existing literature are asymptotic in the sense that the
rate of convergence is attained only after sufficiently many iterations or the initial
guess is sufficiently close to the exact solution. But the convergence estimates we
will present are uniform and they are valid at the very first step of iteration.

The paper is organized as follows. In Section 2 the algorithms are proposed in
a general space decomposition setting. The needed conditions for the convergence
and also the convergence rate analysis are supplied in subsection 2.2. It is shown in
subsection 4.1 that the overlapping domain decomposition is a space decomposition
technique and its convergence does not depend on the mesh size and the number of
subdomains in case a proper coarse mesh is used. The corresponding interpretation
and estimates for multigrid methods is given in subsection 4.2. Applications to the
nonlinear p-Laplace equation are considered in Section 5.

2. AN OPTIMIZATION PROBLEM AND TWO SUBSPACE CORRECTION METHODS

In this section, we shall describe in an abstract fashion a general optimization
problem and two subspace correction iterative methods. Several applications of this
optimization problem can be found in Section 5. Optimal convergence estimates
will be established in the following subsection.

2.1. The optimization problem. Given a reflexive Banach space $V$ and a convex
functional $F : V \rightarrow \mathbb{R}$, we shall consider the nonlinear optimization problem

$$
\min_{v \in V} F(v).
$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
We assume that the functional $F$ is Gateaux differentiable (see [3]) and there exist constants $K, L > 0, p \geq q > 1$ such that
\[
\langle F'(w) - F'(v), w - v \rangle \geq K\|w - v\|_p^p, \quad \forall w, v \in V,
\]
(2)
\[
\|F'(w) - F'(v)\|_{V'} \leq L\|w - v\|_q^{-1}, \quad \forall w, v \in V.
\]
Here $\langle \cdot, \cdot \rangle$ is the duality pairing between $V$ and $V'$ (the dual space of $V$). As a direct consequence of (2), we have
\[
K\|w - v\|_p^p \leq \langle F'(w) - F'(v), w - v \rangle \leq L\|w - v\|_q^{-1}, \forall w, v \in V.
\]
(3)
Under assumption (2), problem (1) has unique solutions (see [12, p. 35]). For some nonlinear problems, the constants $K$ and $L$ may depend on $v$ and $w$.

For simplicity, we set
\[
\sigma = \frac{p}{p - q + 1}, \quad \sigma' = \frac{p}{q - 1}, \quad \text{which satisfy} \quad \frac{1}{\sigma} + \frac{1}{\sigma'} = 1.
\]
Note that $\sigma \leq p$ and the Hölder inequality holds
\[
\sum_{i=1}^{m} |a_i|^{\sigma - 1} |b_i| \leq \left( \sum_{i=1}^{m} |a_i|^p \right)^{\frac{\sigma - 1}{p}} \left( \sum_{i=1}^{m} |b_i|^q \right)^{\frac{q}{p}}.
\]
(4)
The following lemma can be proved in a similar way as [12, p. 25], and the proof can be found in [28].

**Lemma 2.1.** If condition (3) is valid, then
\[
F(w) - F(v) \geq \langle F'(v), w - v \rangle + \frac{K}{p}\|w - v\|_p^p, \quad \forall v, w \in V,
\]
(5)
\[
F(w) - F(v) \leq \langle F'(v), w - v \rangle + \frac{L}{q}\|w - v\|_q^{-1}, \quad \forall v, w \in V.
\]
(6)

We shall use $u$ to denote the unique solution of (1) which satisfies
\[
\langle F'(u), v \rangle = 0, \quad \forall v \in V.
\]
(7)
It is an easy consequence of Lemma 2.1 that
\[
\frac{K}{p}\|v - u\|_p^p \leq F(v) - F(u) \leq \frac{L}{q}\|v - u\|_q^{-1}, \quad \forall v \in V.
\]
(8)

### 2.2. Two subspace correction methods.

We shall now present two iterative methods for solving the optimization problem (1). The methods themselves here are not new and they can be viewed as generalizations of multigrid and domain decomposition methods studied in the literature. As demonstrated in Xu [37], this type of algorithm can be conveniently described and studied in the framework of space decomposition and subspace correction.

Space decomposition refers to a method that decomposes the space $V$ into a sum of closed subspaces, i.e., there are closed subspaces $V_i \subset V, i = 1, 2, \cdots, m$, such that
\[
V = V_1 + V_2 + \cdots + V_m.
\]
(9)
This means that for any $v$, there exists $v_i \in V_i$ such that $v = \sum_{i=1}^{m} v_i$. Following the framework of [37] for linear problems, we consider two types of subspace correction methods based on (9), namely the parallel subspace correction (PSC) method and the successive subspace correction (SSC) method.

The parallel subspace correction method can be described as follows.
Algorithm 2.1. Choose initial value \( u^0 \in V \) and relaxation parameters \( \alpha_i > 0 \) such that \( \sum_{i=1}^{m} \alpha_i \leq 1 \).

1. For \( n \geq 0 \), if \( u^n \in V \) is defined, then find \( e^n_i \in V_i \) in parallel for \( i = 1, 2, \ldots, m \) such that

\[
F(u^n + e^n_i) \leq F(u^n + v_i), \quad \forall v_i \in V_i.
\] (10)

2. Set

\[
u^{n+1} = u^n + \sum_{i=1}^{m} \alpha_i e^n_i,
\] (11) and go to the next iteration.

The successive subspace correction method can be described as follows.

Algorithm 2.2. Choose initial values \( u^0 \in V \).

1. For \( n \geq 0 \), if \( u^n \in V \) is defined, find \( u^{n+i/m} = u^{n+(i-1)/m} + e^n_i \) with \( e^n_i \in V_i \) sequentially for \( i = 1, 2, \cdots, m \) such that

\[
F\left(u^{n+(i-1)/m} + e^n_i\right) \leq F\left(u^{n+(i-1)/m} + v_i\right), \quad \forall v_i \in V_i.
\] (12)

2. Go to the next iteration.

We note that the two algorithms above are well defined since the subspace problems (10) and (12) are uniquely solvable under the assumptions for \( F \) described earlier (see [12]).

For the convergence analysis to be presented in the following section, we shall now introduce two positive constants that in some sense characterize the space decomposition (9).

The first constant, denoted by \( C_1 \), is the least constant satisfying the following property: for any given \( v \in V \), there exist \( v_i \in V_i \) such that

\[
v = \sum_{i=1}^{m} v_i, \quad \text{and} \quad \left( \sum_{i=1}^{m} \|v_i\|_V \right)^{\frac{1}{p}} \leq C_1 \|v\|_V.
\] (13)

The existence of such a constant in an infinitely dimensional Banach space is perhaps not so obvious at first glance, but it can be verified by a simple application of the open mapping theorem.

The second constant, denoted by \( C_2 \), is the least constant satisfying the following property: for any \( w_{ij} \in V, u_i \in V_i \) and \( v_j \in V_j \) the following inequality holds:

\[
\sum_{i,j=1}^{m} \langle F'(w_{ij} + u_i) - F'(w_{ij}), v_j \rangle \leq C_2 \left( \sum_{i=1}^{m} \|u_i\|_V^p \right)^{\frac{1}{p}} \left( \sum_{j=1}^{m} \|v_j\|_V^p \right)^{\frac{1}{p}}.
\] (14)

The existence of \( C_2 \) is obvious by assumption (2). A simple application of the Hölder inequality would give the rough upper bound

\[
C_2 \leq Lm,
\]

but better bounds may be obtained in applications.
Remark 2.1. For some nonlinear problems, the constants $K$ and $L$ may depend on $v$ and $w$. However, we note that Algorithms 1 and 2 are energy decreasing, i.e., $F(u^n) \leq F(u^0)$. Thus, it is easy to prove that there exists a constant $C(u^0)$ which only depends on $u^0$ such that

$$\|u^n\|_V, \|e_i^n\|_V \leq C(u^0), \quad \forall n, i,$$

under the condition that $F$ is coercive, i.e.,

$$F(v) \to \infty \quad \text{as} \quad \|v\|_V \to \infty.$$

Accordingly, one can observe from our analysis given later that we only need assumption (2) on a bounded set $S = \{v|\|v\|_V \leq C(u^0)\}$. It is often true that $K$ and $L$ are uniformly bounded on bounded sets. Therefore, we have stated assumption (2) without explicitly mentioning the dependence of $K$ and $L$ on $u$ and $v$. By omitting this dependence, it offers simplicity and clarity in the analysis.

Domain decomposition methods, multilevel methods and multigrid methods can be viewed as different ways of decomposing finite element spaces into sums of subspaces. See subsections 4.1 and 4.2 for examples of some decompositions of a finite element space and the corresponding estimates for constants $C_1$ and $C_2$. If $F$ is strictly convex, then the iterative solutions of the algorithms converge to the true solution, i.e., they are in a neighbourhood of the true solution. Therefore, we just need to estimate (2) and (14) for $v, w$ and $w_{ij}$ from a neighbourhood of the true solution. For linear problems, estimate (14) is a consequence of the well-known strengthened Cauchy-Schwartz inequality (see Xu [37]).

3. Convergence analysis

Under the assumptions described above, we shall derive in this section uniform convergence rates for both the PSC and SSC iterative algorithms.

Let $u$ be the exact solution of (1) and $u^n$ be the $n$th iterate of algorithm P or S. We need to estimate the rate of reduction of the error $u - u^n$ for each iteration.

To obtain a sharp estimate of this kind, it is important to use an appropriate measurement. For our problem, we find it is convenient to use the measurement

$$d_n = F(u^n) - F(u).$$

Thanks to (8), $d_n$ is more or less like a norm of $u - u^n$.

3.1. Main results. We shall now state the theorems for the convergence rates of Algorithms 2.1 and 2.2. For convenience of exposition, we introduce the parameter

$$r = \frac{p(p-1)}{q(q-1)}.$$

The convergence for Algorithms 2.1 and 2.2 can be stated as:

**Theorem 3.1.** Assume that the space decomposition satisfies (13), (14) and the functional $F$ satisfies (2). Then for both Algorithms 2.1 and 2.2 we have:

1. If $p = q$, then there exists a constant $\delta \in (0, 1)$ depending on $p, q, K, L, C_1$ and $C_2$ (see [32]) such that

$$d_n \leq \delta d_{n-1} \leq \delta^n d_0, \quad \forall n \geq 1.$$
2. If \( p > q \), then there exists a positive constant \( c_0 \) depending on \( d_0, p, q, K, L, C_1 \) and \( C_2 \) (see (33)) such that
\[
d_n \leq \frac{d_{n-1}}{(1 + c_0 d_{n-1}^{(r-1)n})^{1/r}},
\]
(17)
\[
\leq \frac{d_0}{(1 + c_0 d_0^{(r-1)n})^{1/r}}, \quad \forall n \geq 1.
\]

This theorem states that when \( p = q \), the convergence of both algorithms is geometric. In case that \( p > q \), the convergence can be slow, i.e., \( d_n = O\left((rn)^{-\frac{1}{r-1}}\right) \).

Especially, when \( r \) is very big, \( \frac{1}{r} \approx 0 \) and the convergence can be very slow. Using the fact that \( \sigma \leq p \), we see that it is impossible to have \( r < 1 \). In order to have \( r = 1 \), we must require \( p = q \).

3.2. A technical lemma. The proof of our first main result needs the following technical lemma.

Lemma 3.2. Given \( r > 1 \) and \( \eta > 0 \), if \( a \in (0, a_0] \) and \( b > 0 \) satisfy the inequality
\[
b + \eta b^r \leq a,
\]
(18)
then there exists a constant \( \xi_0 = \xi_0(a_0, \eta, r) \in (0, 1) \), depending only on \( a_0, \eta \) and \( r \), such that
\[
b \leq (\eta(r - 1)\xi_0 + a^{1-r})^{1/r} < a.
\]
(19)
Proof. For the given \( a > 0, \eta > 0, r > 1 \), consider the ordinary differential equation
\[
\begin{cases}
\frac{du}{dt} = -\eta u^r, & t > 0, \\
y(0) = a & \text{at } t = 0.
\end{cases}
\]
(20)
Its solution is given by
\[
y(t) = (\eta(r - 1)t + a^{1-r})^{1/r}.
\]
(21)
Consider the function
\[
E(\xi) = y(\xi) + \eta(y(\xi))^r - a.
\]
Note that if \( E(0) = \eta a^r > 0 \) and \( y(\xi) \leq y(0) = a \), we have
\[
E'(\xi) \geq -\eta a^r - \eta^2 r a^{2r-1}
\]
and
\[
E(\xi) = E(0) + \int_0^\xi E'(t)dt \geq \eta a^r - \eta^2 r a^{2r-1} \xi - \eta a^r \xi \geq \eta a^r (1 - \eta a^{r-1} \xi - \xi).
\]
Thus, for $\xi_0 = \frac{1}{\gamma a_0 + 1}$, $E(\xi_0) \geq 0$, namely
\begin{equation}
y(\xi_0) + \eta(y(\xi_0))^r \geq a.
\end{equation}
A combination of (18) and (22) tells that
\begin{align*}
b + \eta b^r & \leq y(\xi_0) + \eta(y(\xi_0))^r.
\end{align*}
By monotonicity of the relevant function, we conclude that
\begin{align*}
b & \leq y(\xi_0).
\end{align*}
This proves the lemma. \hfill \Box

3.3. **Proof of the main theorem.** We are now in a position to present the details of the proof of our main theorem, namely Theorem 3.1.

Using (11), the convexity of $F$ and (5), we deduce that
\begin{align*}
F(u^n) - F(u^{n+1}) &= F(u^n) - F\left(u^n + \sum_{i=1}^{m} \alpha_i e_i^n\right) \\
&= F(u^n) - F\left(\sum_{i=1}^{m} \alpha_i (u^n + e_i^n) + \left(1 - \sum_{i=1}^{m} \alpha_i\right) u^n\right) \\
&\geq F(u^n) - \sum_{i=1}^{m} \alpha_i F(u^n + e_i^n) - \left(1 - \sum_{i=1}^{m} \alpha_i\right) F(u^n) \\
&= \sum_{i=1}^{m} \alpha_i \left(F(u^n) - F(u^n + e_i^n)\right) \\
&\geq - \sum_{i=1}^{m} \alpha_i \|F'(u^n + e_i^n), e_i^n\| + \frac{K}{p} \sum_{i=1}^{m} \alpha_i \|e_i^n\|_V^p \\
&= \frac{K}{p} \sum_{i=1}^{m} \alpha_i \|e_i^n\|_V^p.
\end{align*}

For notational simplicity, we introduce
\begin{align*}
w_j^n &= u^n + \sum_{i=1}^{j} \alpha_i e_i^n.
\end{align*}
It is easy to see that
\begin{equation}
F'(u^n + \sum_{j=1}^{m} \alpha_j e_j^n) - F'(u^n).
\end{equation}
From the property (13) of the space decomposition, there exists $v_i \in V_i$ such that
\begin{equation}
u^{n+1} - u = \sum_{i=1}^{m} v_i \quad \text{and} \quad \left(\sum_{i=1}^{m} \|v_i\|_V^p\right)^{\frac{1}{p}} \leq C_1 \|u^{n+1} - u\|_V.
\end{equation}
We now use (7), (23), (24) and (2) to deduce that

\[
\langle F'(u^{n+1}) - F'(u), u^{n+1} - u \rangle
\]

\[
= \langle F'(u^{n+1}), u^{n+1} - u \rangle = \sum_{i=1}^{m} \langle F'(u^{n+1}), v_i \rangle
\]

(26)

\[
= \sum_{i=1}^{m} \langle F'(u^{n+1}) - F'(u^n + e_i^n), v_i \rangle
\]

\[
= \sum_{i=1}^{m} \langle F'(u^n + \sum_{j=1}^{m} \alpha_j e_j^n) - F'(u^n + e_i^n), v_i \rangle
\]

\[
= \sum_{i=1}^{m} \langle F'(u^n) - F'(u^n + e_i^n), v_i \rangle + \sum_{i=1}^{m} \langle F'(u^n) - F'(u^n + e_i^n), v_i \rangle
\]

\[
= \sum_{i=1}^{m} \sum_{j=1}^{m} \langle F'(w_j^n) - F'(w_j^{n-1}), v_i \rangle
\]

\[
+ \sum_{i=1}^{m} \langle F'(u^n) - F'(u^n + e_i^n), v_i \rangle
\]

\[
\leq C_2 \left( \sum_{j=1}^{m} \| (\alpha_j e_j^n) \|_{\mathcal{V}}^{q-1} \left( \sum_{i=1}^{m} \| v_i \|_{\mathcal{V}}^{q} \right)^{\frac{1}{q}} \right)
\]

\[
+ C_2 \sum_{i=1}^{m} \| e_i^n \|_{\mathcal{V}}^{q-1} \| u \|_{\mathcal{V}}
\]

\[
\leq C_2 \left( \sum_{j=1}^{m} \alpha_j^{\frac{p}{q}} \| e_j^n \|_{\mathcal{V}}^{p} \left( \sum_{i=1}^{m} \| v_i \|_{\mathcal{V}}^{q} \right)^{\frac{1}{q}} \right)
\]

(27)

\[
+ \frac{C_2}{\alpha_{\min}} \left( \sum_{i=1}^{m} \alpha_i \| e_i^n \|_{\mathcal{V}}^{q} \left( \sum_{i=1}^{m} \| v_i \|_{\mathcal{V}}^{q} \right)^{\frac{1}{q}} \right)
\]

\[
\leq C_2 |\alpha_{\max}| \left( \sum_{i=1}^{m} \alpha_i \| e_i^n \|_{\mathcal{V}}^{p} \right)^{\frac{q-1}{p}} \cdot C_1 \| u^{n+1} - u \|_{\mathcal{V}}
\]

(28)

\[
+ C_2 |\min \alpha_i| \cdot \left( \sum_{i=1}^{m} \alpha_i \| e_i^n \|_{\mathcal{V}}^{p} \right)^{\frac{q-1}{p}} \cdot C_1 \| u^{n+1} - u \|_{\mathcal{V}}
\]

\[
\leq C_1 C_2 \left( \frac{\alpha_{\max}^{(p-1)(q-1)}}{p} + \frac{\alpha_{\min}^{(p-1)}}{p} \right)
\]

\[
\cdot \left[ \frac{p}{K} (F(u^n) - F(u^{n+1})) \right]^{\frac{q}{q-1}} \cdot \| u^{n+1} - u \|_{\mathcal{V}}.
\]

In the above, \( \alpha_{\max} \) and \( \alpha_{\min} \) are used to denote

\[
\alpha_{\max} = \max_{1 \leq i \leq m} \alpha_i, \quad \alpha_{\min} = \min_{1 \leq i \leq m} \alpha_i.
\]
By assumption (2) and relation (8), we have
\[
\frac{F'(u^{n+1}) - F'(u)}{\|u^{n+1} - u\|_V} \geq K\|u^{n+1} - u\|_V^{p-1} \geq K \left[ \frac{q}{L} (F(u^{n+1}) - F(u)) \right]^{\frac{p}{q-1}}.
\]

Defining
\[
C^* = \left[ C_1 C_2 \left( \frac{(p-1)(q-1)}{\alpha_{\text{max}}} + \alpha_{\text{min}} \right) \right] \frac{p}{K} \left( \frac{L}{q} \right)^{r}.
\]

one gets from (28) and (29) that
\[
(d_{n+1})^r \leq C^*(d_n - d_{n+1}).
\]

If \( r = 1 \), then from (31) we obtain that
\[
d_{n+1} \leq \delta d_n \quad \text{with} \quad \delta = \frac{C^*}{1 + C^*}.
\]

Next, we consider the case that \( r > 1 \). If \( d_n = 0 \) for an \( n \geq 1 \), then (31) tells us that \( d_k = 0, \forall k \geq n \). In this case, Theorem 3.1 is correct. Now, let us assume that \( d_n > 0, \forall n \geq 1 \). Relation (31) is equivalent to
\[
d_{n+1} + \frac{1}{C^*}(d_{n+1})^r \leq d_n.
\]

An application of Lemma 3.2 ensures that there is an \( \xi_0 = \xi_0(d_0, C^*, r) \in (0, 1) \) such that
\[
d_{n+1} \leq \left( \frac{r - 1}{C^*} + d_{n+1} \right)^{\frac{1}{r}}.
\]

Checking on the value of \( \xi_0 \) from Lemma 3.2 it is easy to see that
\[
c_0 = \frac{(r-1)\xi_0}{C^*} = \frac{r - 1}{rd_{n+1}^{r-1} + C^*}.
\]

By induction, it follows that
\[
d_{n+1} \leq (2c_0 + d_{n+1}^{1-r})^{\frac{1}{1-r}} \leq \cdots \leq (n+1)c_0 + d_n^{1-r})^{\frac{1}{1-r}}.
\]

This proves the theorem for Algorithm 2.1.

Now we proceed with the proof of the theorem for Algorithm 2.2. Notice
\[
F(u^n) - F(u^{n+1}) = \sum_{i=1}^{m} \left[ F(u^{n+i/m}) - F(u^{n+(i-1)/m}) \right].
\]

As \( u^{n+\frac{1}{m}} \) is the minimizer of (10), we get by (5)
\[
F(u^{n+(i-1)/m}) - F(u^{n+i/m}) \geq K \|e_i^n\|_V^{p}.
\]

Thus, estimates (36) and (37) together tell us that
\[
F(u^n) \geq F(u^{n+1})
\]
and

\[ F(u^n) - F(u^{n+1}) \geq \frac{K}{p} \sum_{i=1}^{m} \|e_i^n\|^p V. \]

Similarly to the proofs for (27)–(29), there holds for any \( v_i \in V_i \), which satisfies \( \sum_{i=1}^{m} v_i = u^{n+1} - u \), the relation

\[
\langle F'(u^{n+1}) - F'(u), u^{n+1} - u \rangle = \sum_{i=1}^{m} \langle F'(u^{n+1}) - F'(u^{n+(i-1)/m} + e_i^n), v_i \rangle \]

\[
= \sum_{i=1}^{m} \sum_{j \geq i} \langle F'(u^{n+j/m}) - F'(u^{n+(j-1)/m}), v_i \rangle \]

\[
\leq C_2 \left( \sum_{i=1}^{m} \|e_i^n\|^p V \right)^{\frac{m+1}{p}} \left( \sum_{i=1}^{m} \|v_i\|^p V \right)^{\frac{1}{p}}.
\]

Let \( v_i \) be given as in (25) and using estimates (39) and (40) to obtain

\[
\langle F'(u^{n+1}) - F'(u), u^{n+1} - u \rangle \leq C_2 \left( \sum_{i=1}^{m} \|e_i^n\|^p V \right)^{\frac{m+1}{p}} \left( \sum_{i=1}^{m} \|v_i\|^p V \right)^{\frac{1}{p}}.
\]

\[
\leq C_1 C_2 \left( \sum_{i=1}^{m} \|e_i^n\|^p V \right)^{\frac{m+1}{p}} \cdot \|u^{n+1} - u\| V.
\]

The rest of the proof is the same as the proof for Algorithm 2.1.

Remark 3.1. If there is no extra condition on the decomposed spaces, the condition \( \sum_{i=1}^{m} \alpha_i \leq 1 \) is sufficient and also necessary for the convergence of Algorithm 2.1. In Remark 4.1 of [32, p. 146], an example is given which shows that if \( \sum_{i=1}^{m} \alpha_i > 1 \), then Algorithm 2.1 can be divergent. For overlapping domain decomposition with a suitable coloring, condition \( \sum_{i=1}^{m} \alpha_i \leq 1 \) is nearly optimal. However, for multigrid methods, as we shall discuss later, the upper bound of \( \sum_{i=1}^{m} \alpha_i \) with which the algorithm is convergent can be much larger than 1. The upper bound depends on matrix \( E = (\epsilon_{ij}) \), where \( \epsilon_{ij} \) satisfies

\[
\langle F'(w_{ij} + u_i) - F'(w_{ij}), v_j \rangle \leq \epsilon_{ij} \|u_i\|^q V, \forall w_{ij} \in V_i, \forall u_i \in V_i, \forall v_j \in V_j.
\]

If the decomposed spaces are orthogonal, it is easy to determine the upper bound of \( \sum_{i=1}^{m} \alpha_i \). In computations for general decomposed spaces, a line search to find the value of \( t \), such that the functional

\[ g(t) = F(u^n + t \sum_{i=1}^{m} e_i^n) \]

is attaining its minimum value, would be appropriate. To find such a \( t \), we do not need to solve any system of equations, and we only need to evaluate the functional values, which is not computationally expensive.
4. Space decomposition for $W^{1,p}(\Omega)$

In this section, we consider a special Banach space $V = W^{1,p}(\Omega)$ and then discuss the estimations of the corresponding constant $C_1$ and $C_2$ introduced in Section 2. We first discuss a domain decomposition method and then discuss a multigrid algorithm.

4.1. Overlapping domain decomposition. In this subsection, we show how an overlapping domain decomposition to decompose a finite element space can be used.

Let $\{\Omega_i\}_{i=1}^M$ be a quasi-uniform finite element division or a coarse mesh of $\Omega$, and $\Omega_i$ has diameter of order $H$. For each $\Omega_i$, we further divide it into smaller simplices with diameter of order $h$. In the case that $\Omega$ has a curved boundary, we shall also fill the area between $\partial \Omega$ and $\partial \Omega_H$; here $\Omega_H = \bigcup_{i=1}^M \Omega_i$ with finite elements with diameters of order $h$. We assume that the resulting elements form a shape regular finite element subdivision of $\Omega$ (see Ciarlet [7]). We call this the fine mesh or the $h$-level subdivision of $\Omega$ (see Ciarlet [7]). We denote the piecewise $r$th order polynomial finite element spaces with zero trace on $\partial \Omega_H$ and $\partial \Omega_i$, over the $H$-level and $h$-level subdivisions of $\Omega$, respectively. More specifically,

$$S_0^H = \left\{ v \in W_0^{1,\infty}(\Omega_H) \mid v|_{\Omega_i} \in P_r(\Omega_i), \forall i \right\},$$

$$S_0^h = \left\{ v \in W_0^{1,\infty}(\Omega_h) \mid v|_{\mathcal{T}} \in P_r(\mathcal{T}), \forall \mathcal{T} \in \mathcal{T}_h \right\},$$

For each $\Omega_i$, we consider an enlarged subdomain $\Omega_i^\delta$ consisting of elements $\mathcal{T}, \mathcal{T} \in \mathcal{T}_h$ with $\text{dist}(\mathcal{T}, \Omega_i) \leq \delta$. The union of $\Omega_i^\delta$ covers $\Omega_h$ with overlaps of size $\delta$. Let us denote the piecewise $r$th order polynomial finite element space with zero traces on the boundaries $\partial \Omega_H$ as $S_0^0(\Omega_i^\delta)$. Then one can show that

$$S_0^h = S_0^H + \sum_{i=1}^M S_0^h(\Omega_i^\delta). \tag{42}$$

For the overlapping subdomains, assume that there exist $m$ colors such that each subdomain $\Omega_i^\delta$ can be marked with one color, and the subdomains with the same color will not intersect with each other. For suitable overlaps, one can always choose $m = 2$ if $d = 1$; $m \leq 4$ if $d = 2$; $m \leq 8$ if $d = 3$ (see Figure 1). Let $\Omega_i^\delta$ be the union of the subdomains with the $i$th color and

$$V_i = \{ v \in S_0^h \mid v(x) = 0, \ x \notin \Omega_i^\delta \}.$$  

By denoting subspaces $V_0 = S_0^H, V = S_0^h$, we find that decomposition (42) means

$$V = V_0 + \sum_{i=1}^m V_i, \tag{43}$$

and so the two level method is a way to decompose the finite element space. Following an argument in [35], let $\{\theta_i\}_{i=1}^m$ be a partition of unity with respect to $\{\Omega_i^\delta\}_{i=1}^m$, i.e., $\theta_i \in C_0^\infty(\Omega_i^\delta \cap \Omega), \theta_i \geq 0$ and $\sum_{i=1}^m \theta_i = 1$. It can be chosen so that

$$|\nabla \theta_i| \leq C/\delta, \quad \theta_i(x) = \begin{cases} 1 & \text{if dist}(x, \partial \Omega_i^\delta) \geq \delta \text{ and } x \in \Omega_i^\delta, \\ 0 & \text{on } \Omega \cap \partial \Omega_i^\delta. \end{cases}$$

Let $I_h$ be an interpolation operator which uses the function values at the $h$-level nodes. For any $v \in V$, let $v_0 \in V_0$ be the $L^2$ projection of $v$, namely, $(v_0, \phi_H) = (v, \phi_H), \forall \phi_H \in V_0$, and $v_i = I_h(\theta_i(v - v_0))$. They satisfy $v = \sum_{i=0}^m v_i$, and
Lemma 4.1. For any $s \geq 1$,

$$
\left( \|v_0\|^s_{1,p} + \sum_{i=1}^{m} \|v_i\|^s_{1,p} \right)^{\frac{1}{s}} \leq C(m+1)^{\frac{1}{s}} \left( 1 + \left( \frac{H}{\delta} \right)^{\frac{s}{s-p}} \right) \|v\|_{1,p}.
$$

The proof of the lemma above is essentially similar to the well-known case of $s = p = 2$, except we need to use the following $W^{1,p}$ stability estimate for $L^2$ projection:

$$
\int_{\Omega} |\nabla v_0|^p dx dy \leq C \int_{\Omega} |\nabla v|^p dx dy.
$$

This stability result is easy to prove by using, for example, the local $L^2$-projection technique in [36] (for details, see [34]). Estimate (44) shows that for overlapping domain decomposition, the constants in (13) and (14) are

$$
C_1 = C(m) \left( 1 + \left( \frac{H}{\delta} \right)^{\frac{s}{s-p}} \right), \quad C_2 = C_{\delta} m.
$$

By requiring $\delta = c_0 H$, where $c_0$ is a given constant, we have that $C_1$ and $C_2$ are independent of the mesh parameters $h$ and $H$, the number of subdomains. So if the proposed algorithms are used, their error reductions per step are independent of these parameters.

4.2. Multigrid decomposition. In this subsection, we discuss the application of our theory to multigrid methods. From the space decomposition point of view,
a multigrid algorithm is built upon the subspaces that are defined on a nested sequence of finite element partitions.

We assume that the finite element partition $T$ is constructed by a successive refinement process. More precisely, $T = T_J$ for some $J > 1$, and $T_j$ for $j \leq J$ are a nested sequence of quasi-uniform finite element partitions, i.e., $T_j$ consist of finite elements $T_j = \{ \tau_j^i \}$ of size $h_j$ such that $\Omega = \bigcup \tau_j^i$ for which the quasi-uniformity constants are independent of $j$ (cf. [7]), and $\tau_j^i$ is a union of elements of $\{ \tau_j^i \}$. We further assume that there is a constant $\gamma < 1$, independent of $j$, such that $h_j$ is proportional to $2^{-j}$.

As an example, in the two dimensional case, a finer grid is obtained by connecting the midpoints of the edges of the triangles of the coarser grid, with $T_1$ being the given coarsest initial triangulation, which is quasi-uniform. In this example, $\gamma = 1/\sqrt{2}$. We can use much smaller $h_j$ in constructing the meshes, but the constant $C_1$ is getting larger when $\gamma$ is becoming smaller (see (46)).

Corresponding to each finite element partition $T_j$, a finite element space $M_j$ can be defined by

$$ M_j = \{ v \in W_1^1(\Omega) : v|_\tau \in P_1(\tau), \quad \forall \tau \in T_j \}. $$

Each finite element space $M_j$ is associated with a nodal basis, denoted by $\{ \phi_j^i \}_{i=1}^{n_j}$ satisfying

$$ \phi_j^i(x_j^k) = \delta_{ik}, $$

where $\{ x_j^k \}_{k=1}^{n_j}$ is the set of all nodes of the elements of $T_j$. Associated with each such nodal basis function, we define a one dimensional subspace as

$$ M_j^i = \text{span} \{ \phi_j^i \}. $$

On each level, the nodes can be colored so that the neighboring nodes are always of different colors. The number of colors needed for a regular mesh is always a bounded constant; call it $m_c$. Let $V_j^k$, $k = 1, 2, \ldots, m_c$ be the sum of the subspaces $M_j^i$ associated with nodes of the $k$th color on level $j$. Letting $V = M_J$, we have the following trivial space decomposition:

$$ V = \sum_{j=1}^J \sum_{k=1}^{m_c} V_j^k. $$

(45)

Each subspace $V_j^k$ contains some orthogonal one dimensional subspaces $M_j^i$, and so the minimization problems (10) and (12) for each $V_j^k$ can be done in parallel over the one dimensional subspaces $M_j^i$.

4.2.1. Estimation of the constant $C_1$. For any $j \leq J$, let $Q_j$ be the $L^2$ project operator to the finite element space $M_j$ at level $j$. For any $v \in V$, define $v_j = (Q_j - Q_{j-1})v \in M_j$. A further decomposition of $v_j$ is given by

$$ v_j = \sum_{i=1}^{n_j} v_j^i \quad \text{with} \quad v_j^i = v_j(x_j^i)\phi_j^i. $$

Let $v_j^k$, $k = 1, 2, \ldots, m_c$ be the sum of $v_j^i$ associated with the nodes of the $k$th color on level $j$. It is easy to see that

$$ v_j = \sum_{k=1}^{m_c} v_j^k = \sum_{i=1}^{n_j} v_j^i. $$
Denote $\Omega^k_j$ the union of the support sets of the basis functions associated with the $k$th color nodes on level $j$. We estimate
\[
\sum_{k=1}^{m_c} |v_j^k|_{1,p}^\sigma = \sum_{k=1}^{m_c} \left( \sum_{x^k_j \in \Omega^k_j} |v_j(x^k_j)|^p \phi_j^k \right)^{\frac{\sigma}{p}} \leq C h_j^{-\sigma} \sum_{k=1}^{m_c} \left( \sum_{x^k_j \in \Omega^k_j} |v_j(x^k_j)|^p \right)^{\frac{\sigma}{p}}.
\]

In the above, we have assumed that $\Omega \subset R^d$, $d = 1, 2, 3, \ldots$. Using the inequality
\[
\sum_{k=1}^{m_c} |a_k|^\sigma \leq \left( \sum_{k=1}^{m_c} |a_k|^p \right)^{\frac{\sigma}{p}} \left( m_c \right)^{1-\frac{\sigma}{p}},
\]
we get that
\[
\sum_{k=1}^{m_c} |v_j^k|_{1,p}^\sigma \leq C h_j^{-\sigma} \left( m_c \right)^{1-\frac{\sigma}{p}} \left( \sum_{i=1}^{n_j} |v_j(x^k_i)|^p \right)^{\frac{\sigma}{p}} \leq C h_j^{-\sigma} \|v_j\|_{0,p}^\sigma.
\]

Here, we have used the fact that, in the finite element space, an $L^p$ norm is equivalent to some discrete $L^p$ norm, namely
\[
\|v_j\|_{0,p}^p \cong h_j^d \sum_{i=1}^{n_j} |v_j(x^k_i)|^p.
\]

As a consequence,
\[
\sum_{j=1}^J \sum_{k=1}^{m_c} \|v_j^k\|_{1,p}^\sigma \leq C \sum_{j=1}^J h_j^{-\sigma} \|v_j\|_{0,p}^\sigma \leq C \sum_{j=1}^J h_j^{-\sigma} \|Q_j - Q_{j-1}\| \leq C \sum_{j=1}^J h_j^{-\sigma} \|I - Q_{j-1}\| \leq C \gamma^{-2\sigma} J \|v\|_{1,p}^\sigma,
\]
which proves that
\[
C_1 \cong J^\frac{1}{2} \cong |\log h_j|^\frac{1}{2}.
\]

In proving inequality (46), we have used the stability in $L^p$ of the $L^2$-projection (9) and the error estimate for $L^2$-projections (see [7]).

4.2.2. Estimation of the constant $C_2$. From condition (5), we see that
\[
\langle F'(w_{ij} + u_i) - F'(w_{ij}), v_j \rangle \leq L \|u_i\|_{V} \|v_j\|_{V}.
\]

However, in order to estimate the constant $C_2$, we need to use a finer estimate than (47). For any $w, u, v \in V$, we need the functional $F$ to satisfy
\[
\langle F'(w + u) - F'(w), v \rangle \leq L \|u\|_{1,p, \supp(u) \cap \supp(v)} \|v\|_{1,p, \supp(u) \cap \supp(v)}.
\]

In the above and also later, $\supp$ denotes the support set of a function. For any $u \in M^k_j$ and $v \in M^k_l$, $j \leq l$, we note that the size of $\supp(u) \cap \supp(v)$ is at most the size of $\supp(v)$. Thus since $u$ is piecewise linear,
\[
\|u\|_{1,p, \supp(u) \cap \supp(v)} \leq C \gamma^{\frac{d}{p} |j-l|} \|u\|_{1,p}, \quad \forall u \in M^k_j, \forall v \in M^k_l.
\]
Let \( w \in V, u \in V^j_i \) and \( v \in V^k_l \). We decompose \( u \) and \( v \) as

\[
u = \sum_{\alpha=1}^{n_j} u_{\alpha}, \quad u_{\alpha} = u(x_\alpha^\sigma) \phi_j^\alpha, \quad v = \sum_{\beta=1}^{n_k} v_{\beta}, \quad v_{\beta} = v(x_\beta^\sigma) \phi_i^\beta,
\]
i.e., functions \( u \) and \( v \) are decomposed into functions from the one dimensional subspaces of the same colors. We shall assume that the following inequality is valid for the above decomposition:

\[
\langle F'(w + \sum u_{\alpha}) - F'(w), \sum v_{\beta} \rangle \leq \sum_{\alpha} \sum_{\beta} \langle F'(w + u_{\alpha}) - F'(w), v_{\beta} \rangle.
\]

The above inequality is often a consequence of the orthogonality of the one dimensional subspaces of the same color and the fact that \( u \) is zero at the nodes that do have the color of \( u \).

From [50], [68], [69] and the orthogonality of the one dimensional subspaces of the same color, it is easy to see that

\[
\langle F'(w + u) - F'(w), v \rangle \leq C\gamma^{\frac{2d(q-1)}{p}(l-j)} L \sum_{\alpha} \sum_{\beta} \|u_{\alpha}\|_{1,p}^{q-1} \|v_{\beta}\|_{1,p}
\]

\[
= C\gamma^{\frac{2d(q-1)}{p}(l-j)} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V^j_i, \ v \in V^k_l, \ j \leq l.
\]

For \( j > l \), we shall have

\[
\langle F'(w + u) - F'(w), v \rangle \leq C\gamma^{\frac{2d(q-1)}{p}(l-j)} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V^j_i, \ v \in V^k_l.
\]

Denoting \( \gamma_0 = \gamma^{\frac{2d(q-1)}{p} \max(l-1,1)} \), we get from the above two estimates

\[
\langle F'(w + u) - F'(w), v \rangle \leq C\gamma_0^{\frac{2d(q-1)}{p}(l-j)} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V^j_i, \ v \in V^k_l.
\]

To estimate the constant \( C_2 \) we need the next lemma, which extends a result of [27] p.184].

**Lemma 4.2.** Let \( A = (A_{ij} \theta_{ij}) \) be an \( n_1 \times n_2 \) matrix. Then

\[
\|Ax\|_{\sigma} \leq \max_i \left( \sum_j |A_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \max_j \left( \sum_i |A_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \|x\|_{\sigma}.
\]

**Lemma 4.3** (Proof of Lemma 4.2). The Cauchy-Schwarz inequality gives

\[
\|Ax\|_{\sigma} = \sum_i \left| \sum_j A_{ij} \theta_{ij} x_j \right|^{\sigma}
\]

\[
\leq \sum_i \left( \sum_j |A_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \left( \sum_j |\theta_{ij}|^{\sigma} |x_j|^{\sigma} \right)
\]

\[
\leq \max_i \left( \sum_j |A_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \sum_j \left( \sum_i |\theta_{ij}|^{\sigma} |x_j|^{\sigma} \right)
\]

\[
\leq \max_j \left( \sum_i |\theta_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \max_i \left( \sum_j |A_{ij}|^{\sigma} \right)^{\frac{1}{\sigma}} \sum_j |x_j|^{\sigma},
\]

which proves the lemma. \( \square \)
Using the above inequality and Corollary 4.1, we get that
\[ \|Ax\|_{\ell^p} \leq \left( \max_i \sum_j |A_{ij}| \right) \|x\|_{\ell^p}. \]

**Proof of Corollary 4.1.** It is easy to see that
\[ |A_{ij}| = |A_{ij}|^{\frac{1}{p}} |A_{ij}|^{\frac{1}{q}}. \]
The Corollary is an easy consequence of Lemma 4.2 by setting \( A_{ij} := |A_{ij}|^{\frac{1}{p}} \) and \( \theta_{ij} := |A_{ij}|^{\frac{1}{q}} \).

For a given \( u_j^i \in V_j^i \) and \( v^k \in V_t^k \), an application of (4) gives
\[
\sum_{i=1}^{m_e} \sum_{k=1}^{m_e} \|u_j^i\|_{1,p}^{q-1} \|v^k\|_{1,p} \leq m_e \left( \sum_{i=1}^{m_e} \|u_j^i\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \left( \sum_{k=1}^{m_e} \|v^k\|_{1,p}^{p} \right)^{\frac{q-1}{p}}.
\]

Using the above inequality and Corollary 4.1, we get that
\[
\sum_{j=1}^{J} \sum_{i=1}^{J} \sum_{i=1}^{m_e} \sum_{k=1}^{m_e} \gamma_0^{l-1} \|u_j^i\|_{1,p} \|v^k\|_{1,p} \leq C_{m_e} \left( \max_j \sum_{i=1}^{J} \gamma_0^{l-1} \|u_j^i\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \left( \sum_{k=1}^{m_e} \|v^k\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \leq C_{m_e} \left( \sum_{j=1}^{J} \sum_{i=1}^{m_e} \|u_j^i\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \left( \sum_{k=1}^{m_e} \|v^k\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \leq C_{m_e} \frac{1}{1-\gamma_0} \left( \sum_{j=1}^{J} \sum_{i=1}^{m_e} \|u_j^i\|_{1,p}^{p} \right)^{\frac{q-1}{p}} \left( \sum_{k=1}^{m_e} \|v^k\|_{1,p}^{p} \right)^{\frac{q-1}{p}}.
\]

From (14), (51) and (52), we conclude that the constant \( C_2 \) is independent of the mesh size \( h \) and the number of levels \( J \) for decomposition (14).

**Remark 4.1.** In the case that \( p = q \), the estimations we have derived for the constants \( C_1 \) and \( C_2 \) are also valid for the decomposition
\[
V = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \mathcal{M}_j^i,
\]
i.e., the coloring is not necessary for implementing the algorithms.

**Remark 4.2.** We would like to note that the multigrid algorithm described in this section is not quite optimal and it needs \( O(n_J \log n_J) \) operations in each iteration.

5. An Application

In this section we give some problems to which our algorithms and theory can be applied without going into the details of analysis. Other applications are also possible: for example, the eigenvalue problem in Chan and Sharapov [3, 26] and also some other nonlinear partial differential equations for minimum surface, superconductivity and porous media flows.
We consider the following nonlinear problem:

\begin{equation}
\begin{cases}
- \nabla \cdot ([|\nabla u|^{s-2}\nabla u]_{s}) = f \text{ in } \Omega \subset \mathbb{R}^d \ (1 < s < \infty), \\
u = 0 \text{ on } \partial \Omega.
\end{cases}
\end{equation}

For equation (54), we assume \( f \in W^{-1,s'}(\Omega), \frac{1}{s} + \frac{1}{s'} = 1 \). By standard techniques, it can be shown (see [12]), that (54) possesses a unique solution which is the minimizer of

\[ \min_{v \in W^{1,s}_0(\Omega)} \left( \frac{1}{s} \int_{\Omega} |\nabla v|^s - \langle f, v \rangle \right). \]

Even with very smooth data, the solution \( u \) may not be in the space \( W^{2,s}_0(\Omega) \) (see Ciarlet [7, p. 324]). When \( s \) is close to 1 or is very big \( (s \gg 2) \), it is difficult to solve this problem numerically. It can proven that conditions (2) are valid for this problem (see p. 319 and p. 325 of Ciarlet [7]). More precisely, we have for

\[ V = W^{1,s}_0(\Omega), \quad F(v) = \int_{\Omega} \left( \frac{1}{s} |\nabla v|^s - f v \right) dx \]

the following estimates:

\begin{align*}
(55) & \quad \langle F'(v) - F'(w), v - w \rangle \geq \|v - w\|_{1,s}^s, \quad \text{if } s \geq 2; \\
(56) & \quad \langle F'(v) - F'(w), v - w \rangle \geq \alpha \|v - w\|_{1,s}^2, \quad \text{if } 1 < s \leq 2; \\
(57) & \quad \|F'(v) - F'(w)\|_{V'} \leq \beta \|v - w\|_{1,s}^{s-2}, \quad \text{if } s \geq 2; \\
(58) & \quad \|F'(v) - F'(w)\|_{V'} \leq \beta \|v - w\|_{1,s}^{s-1}, \quad \text{if } 1 < s \leq 2.
\end{align*}

In the above, \( \alpha \) and \( \beta \) are independent of \( v \) and \( w \) and are strictly positive. The proof of (55) and (57) is given in p. 319 of Ciarlet [7]. The proof of (56) and (58) can be found in Glowinski and Marrocco [14]. Corresponding to condition (2), these estimates imply that

\[ p = s, \quad q = 2 \quad \text{if } s \geq 2; \]

\[ p = 2, \quad q = s \quad \text{if } 1 < s \leq 2. \]

It is apparent that the functional \( F \) is coercive:

\[ F(v) \geq \frac{1}{2s} |v|_{1,s}^s - C_3 \|f\|_{s,s'} \]

for some constant \( C_3 \). From this and (55)-(58), it is easy to see that \( K \) and \( L \) are uniformly bounded on the bounded set \( S = \{v \ | \ |v|_V \leq C(u^0)\} \). Choosing \( C(u^0) \) large enough, we can see that all the \( u \) and \( v \), where we have applied assumption (2), are in fact in \( S \).

An application of Theorem 3.1 gives the rate of convergence, i.e.,

\[ d_n \leq (Cn)^{-\frac{2}{s-2}}, \quad \text{for } s > 2; \]

\[ d_n \leq \delta^n d_0, \quad \text{for } s = 2; \]

\[ d_n \leq (Cn)^{-\frac{(s-1)}{(s-1)}} \quad \text{for } 1 < s < 2. \]

The constant \( C \) only depends on \( d_0, K, L, C_1, C_2 \) and \( s \).
We would like to point out that similar results can be obtained for a more general problem

\[
\min_{v \in W_0^{1,\infty}(\Omega)} \int_{\Omega} \frac{1}{2} a(|\nabla v|^2) + f(v).
\]

We assume that \(a\) is strictly convex and \(f\) is convex and both are differentiable.

If \(V_i\) are the domain decomposition subspaces, then corresponding subspace problem (10) or (12) is a nonlinear problem in each subdomain, which has a smaller size than the original problem. For some minimization methods, convergence and computing time depend on the size of the problem. Thus by first reducing the problem into smaller size problems and then minimizing, we may gain efficiency. If \(V_i\) are the multigrid nodal basis subspaces, then the subspace problem is a one dimensional nonlinear problem and we can use efficient minimization routines to solve the one dimensional problems.

Acknowledgments. The authors would like to thank Steinar Evje for valuable discussions related to the proofs of Lemma 3.2 Professor M. Espedal for some earlier participation in this work, and Professor A. Zhou for discussions and help.

References


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BERGEN, JOHANNES BRUNSGATE 12, 5007, BERGEN, NORWAY

E-mail address: tai@mi.uib.no
URL: http://www.mi.uib.no/~tai

CENTER FOR COMPUTATIONAL MATHEMATICS AND APPLICATIONS AND DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PENNSYLVANIA 16802

E-mail address: xu@math.psu.edu
URL: http://www.math.psu.edu/xu/