MATHEMATICS OF COMPUTATION
Volume 71, Number 237, Pages 105–124
S 0025-5718(01)01311-4
Article electronically published on May 11, 2001

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

XUE–CHENG 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], Deuflhard 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

Received by the editor March 10, 1998 and, in revised form, September 15, 1999 and March 24, 2000.

2000 *Mathematics Subject Classification.* Primary 65N22, 65N55, 65K10, 65J15.

Key words and phrases. Parallel, multigrid, domain decomposition, nonlinear, elliptic equation, space decomposition, convex optimization.

The work of the first author was partially supported by the Norwegian Research Council Strategic Institute Program within Inverse Problems at RF–Rogaland Research.

The work of the second author was partially supported by NSF DMS-9706949 NSF ACI-9800244, NASA NAG2-1236 through Pennsylvania State University.

©2001 American Mathematical Society

[24], Dryja and Hackbusch [10], Kornhuber [17, 18], Tai and Espedal [33], and Zou [40].

Our algorithms bear some of the natures of the methods of Mandel [20], Gelman 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] differers 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 problems which can handle some nonlinear diffusion problems even when the nonlinear diffusion coefficient is degenerate or singular (see Section 5).

The iterative methods we will study in this paper can be viewed as a straightforward 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 techniques 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 $F : V \mapsto R$, we shall consider the nonlinear optimization problem

$$\min_{v \in V} F(v). \tag{1}$$

We assume that the functional F is Gateaux differentiable (see [5]) and there exist constants $K,\ L > 0$, $p \geq q > 1$ such that

$$\begin{aligned} \langle F'(w) - F'(v), w - v\rangle &\geq K\|w - v\|_V^p, \quad \forall w, v \in V, \\ \|F'(w) - F'(v)\|_{V'} &\leq L\|w - v\|_V^{q-1}, \quad \forall w, v \in V. \end{aligned} \tag{2}$$

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

$$K\|w - v\|_V^p \leq \langle F'(w) - F'(v), w - v\rangle \leq L\|w - v\|_V^q, \forall w, v \in V. \tag{3}$$

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

$$\sum_{i=1}^{m} |a_i|^{q-1}|b_i| \leq \left(\sum_{i=1}^{m} |a_i|^p\right)^{\frac{q-1}{p}} \left(\sum_{i=1}^{m} |b_i|^\sigma\right)^{\frac{1}{\sigma}}. \tag{4}$$

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*

$$F(w) - F(v) \geq \langle F'(v), w - v\rangle + \frac{K}{p}\|w - v\|_V^p, \quad \forall v, w \in V, \tag{5}$$

$$F(w) - F(v) \leq \langle F'(v), w - v\rangle + \frac{L}{q}\|w - v\|_V^q, \quad \forall v, w \in V. \tag{6}$$

We shall use u to denote the unique solution of (1) which satisfies

$$\langle F'(u), v\rangle = 0, \quad \forall v \in V. \tag{7}$$

It is an easy consequence of Lemma 2.1 that

$$\frac{K}{p}\|v - u\|_V^p \leq F(v) - F(u) \leq \frac{L}{q}\|v - u\|_V^q, \quad \forall v \in V. \tag{8}$$

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

$$V = V_1 + V_2 + \cdots + V_m. \tag{9}$$

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 $e_i^n \in V_i$ in parallel for $i = 1, 2, \cdots, m$ such that*

$$F\left(u^n + e_i^n\right) \leq F\left(u^n + v_i\right), \quad \forall v_i \in V_i. \tag{10}$$

2. *Set*

$$u^{n+1} = u^n + \sum_{i=1}^m \alpha_i e_i^n, \tag{11}$$

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 $u^{n+i/m} = u^{n+(i-1)/m} + e_i^n$ with $e_i^n \in V_i$ sequentially for $i = 1, 2, \cdots, m$ such that*

$$F\left(u^{n+(i-1)/m} + e_i^n\right) \leq F\left(u^{n+(i-1)/m} + v_i\right), \quad \forall v_i \in V_i. \tag{12}$$

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

$$v = \sum_{i=1}^m v_i, \qquad \text{and} \qquad \left(\sum_{i=1}^m \|v_i\|_V^\sigma\right)^{\frac{1}{\sigma}} \leq C_1 \|v\|_V. \tag{13}$$

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:

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

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., $F(u^n) \le F(u^0)$. Thus, it is easy to prove that there exists a constant $C(u^0)$ which only depends on u^0 such that

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

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 $S = \{v | \, \|v\|_V \le C(u^0)\}$. 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 nth 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

$$d_n = F(u^n) - F(u). \tag{15}$$

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*

$$d_n \le \delta d_{n-1} \le \delta^n d_0, \quad \forall n \ge 1. \tag{16}$$

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*

$$\begin{aligned} d_n &\leq \frac{d_{n-1}}{(1+c_0 d_{n-1}^{r-1})^{\frac{1}{r-1}}} \\ &\leq \frac{d_0}{(1+c_0 d_0^{r-1} n)^{\frac{1}{r-1}}}, \quad \forall n \geq 1. \end{aligned} \tag{17}$$

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, $\frac{1}{1-r} \approx 0$ 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*

$$b + \eta b^r \leq a, \tag{18}$$

then there exists a constant $\xi_0 = \xi_0(a_0, \eta, r) \in (0, 1)$, depending only on a_0, η and r, such that

$$b \leq \left(\eta(r-1)\xi_0 + a^{1-r}\right)^{\frac{1}{1-r}} < a. \tag{19}$$

Proof. For the given $a > 0, \eta > 0, r > 1$, consider the ordinary differential equation

$$\begin{cases} \frac{dy}{dt} &= -\eta y^r, \quad t > 0, \\ y &= a \quad \text{at} \quad t = 0. \end{cases} \tag{20}$$

Its solution is given by

$$y(t) = \left(\eta(r-1)t + a^{1-r}\right)^{\frac{1}{1-r}}. \tag{21}$$

Consider the function

$$E(\xi) \equiv y(\xi) + \eta(y(\xi))^r - a.$$

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

$$E(\xi) = E(0) + \int_0^\xi E'(t)dt \geq \eta a^r - \eta^2 r a^{2r-1}\xi - \eta a^r \xi \geq \eta a^r (1 - \eta r a_0^{r-1}\xi - \xi).$$

Thus, for $\xi_0 = \frac{1}{\eta r a_0^{r-1}+1}$, $E(\xi_0) \geq 0$, namely

$$y(\xi_0) + \eta(y(\xi_0))^r \geq a. \tag{22}$$

A combination of (18) and (22) tells that

$$b + \eta b^r \leq y(\xi_0) + \eta(y(\xi_0))^r.$$

By monotonicity of the relevant function, we conclude that

$$b \leq y(\xi_0).$$

This proves the lemma. □

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

(23)

$$\begin{aligned}
F(u^n) - F(u^{n+1}) &= F(u^n) - F\Big(u^n + \sum_{i=1}^m \alpha_i e_i^n\Big) \\
&= F(u^n) - F\Big(\sum_{i=1}^m \alpha_i (u^n + e_i^n) + \Big(1 - \sum_{i=1}^m \alpha_i\Big) u^n\Big) \\
&\geq F(u^n) - \sum_{i=1}^m \alpha_i F\left(u^n + e_i^n\right) - \Big(1 - \sum_{i=1}^m \alpha_i\Big) F(u^n) \\
&= \sum_{i=1}^m \alpha_i \Big(F(u^n) - F(u^n + e_i^n)\Big) \\
&\geq -\sum_{i=1}^m \alpha_i \langle F'(u^n + e_i^n), e_i^n\rangle + \frac{K}{p}\sum_{i=1}^m \alpha_i \|e_i^n\|_V^p \\
&= \frac{K}{p}\sum_{i=1}^m \alpha_i \|e_i^n\|_V^p.
\end{aligned}$$

For notational simplicity, we introduce

$$w_j^n = u^n + \sum_{i=1}^j \alpha_i e_i^n.$$

It is easy to see that

$$F'\Big(u^n + \sum_{j=1}^m \alpha_j e_j^n\Big) - F'(u^n). \tag{24}$$

From the property (13) of the space decomposition, there exists $v_i \in V_i$ such that

$$u^{n+1} - u = \sum_{i=1}^m v_i \quad \text{and} \quad \left(\sum_{i=1}^m \|v_i\|_V^\sigma\right)^{\frac{1}{\sigma}} \leq C_1 \|u^{n+1} - u\|_V. \tag{25}$$

We now use (7), (23), (24) and (2) to deduce that

$$
\begin{aligned}
&\left\langle F'(u^{n+1}) - F'(u), u^{n+1} - u \right\rangle \\
&= \langle F'(u^{n+1}), u^{n+1} - u \rangle = \sum_{i=1}^{m} \left\langle F'(u^{n+1}), v_i \right\rangle \\
&= \sum_{i=1}^{m} \left\langle F'(u^{n+1}) - F'(u^n + e_i^n), v_i \right\rangle \qquad (26)\\
&= \sum_{i=1}^{m} \left\langle F'\Big(u^n + \sum_{j=1}^{m} \alpha_j e_j^n\Big) - F'(u^n + e_i^n), v_i \right\rangle \\
&= \sum_{i=1}^{m} \left\langle F'\Big(u^n + \sum_{j=1}^{m} \alpha_j e_j^n\Big) - F'(u^n), v_i \right\rangle \\
&\quad + \sum_{i=1}^{m} \left\langle F'(u^n) - F'(u^n + e_i^n), v_i \right\rangle \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} \left\langle F'(w_j^n) - F'(w_{j-1}^n), v_i \right\rangle \\
&\quad + \sum_{i=1}^{m} \left\langle F'(u^n) - F'(u^n + e_i^n), v_i \right\rangle \\
&\le C_2 \Big(\sum_{j=1}^{m} \|(\alpha_j e_j^n)\|_V^p \Big)^{\frac{q-1}{p}} \Big(\sum_{i=1}^{m} \|v_i\|_V^\sigma \Big)^{\frac{1}{\sigma}} \\
&\quad + C_2 \sum_{i=1}^{m} \|e_i^n\|_V^{q-1} \|v_i\|_V \\
&\le C_2 \Big(\sum_{j=1}^{m} \alpha_j^p \|e_j^n\|_V^p \Big)^{\frac{q-1}{p}} \Big(\sum_{i=1}^{m} \|v_i\|_V^\sigma \Big)^{\frac{1}{\sigma}} \\
&\quad + \frac{C_2}{\alpha_{min}^{\frac{q-1}{p}}} \Big(\sum_{i=1}^{m} \alpha_i \|e_i^n\|_V^p \Big)^{\frac{q-1}{p}} \Big(\sum_{i=1}^{m} \|v_i\|_V^\sigma \Big)^{\frac{1}{\sigma}} \qquad (27)\\
&\le C_2 |\alpha_{max}|^{\frac{(p-1)(q-1)}{p}} \Big(\sum_{i=1}^{m} \alpha_i \|e_i^n\|_V^p \Big)^{\frac{q-1}{p}} \cdot C_1 \|u^{n+1} - u\|_V \\
&\quad + C_2 |\min \alpha_i|^{-\frac{q-1}{p}} \Big(\sum_{i=1}^{m} \alpha_i \|e_i^n\|_V^p \Big)^{\frac{q-1}{p}} \cdot C_1 \ \|u^{n+1} - u\|_V \qquad (28)\\
&\le C_1 C_2 \Big(\alpha_{max}^{\frac{(p-1)(q-1)}{p}} + \alpha_{min}^{-\frac{q-1}{p}} \Big) \\
&\quad \cdot \Big[\frac{p}{K} \big(F(u^n) - F(u^{n+1}) \big) \Big]^{\frac{q-1}{p}} \cdot \|u^{n+1} - u\|_V .
\end{aligned}
$$

In the above, α_{max} and α_{min} are used to denote

$$\alpha_{max} = \max_{1 \le i \le m} \alpha_i, \qquad \alpha_{min} = \min_{1 \le i \le m} \alpha_i .$$

By assumption (2) and relation (8), we have

$$\begin{aligned} &\frac{\langle F'(u^{n+1}) - F'(u), u^{n+1} - u\rangle}{\|u^{n+1} - u\|_V} \\ &\geq \quad K\|u^{n+1} - u\|_V^{p-1} \geq K\left[\frac{q}{L}(F(u^{n+1}) - F(u))\right]^{\frac{p-1}{q}}. \end{aligned} \tag{29}$$

Defining

$$C^* = \left[\frac{C_1C_2\left(\alpha_{max}^{\frac{(p-1)(q-1)}{p}} + \alpha_{min}^{-\frac{q-1}{p}}\right)}{K}\right]^{\frac{p}{q-1}} \frac{p}{K}\left(\frac{L}{q}\right)^r, \tag{30}$$

one gets from (28) and (29) that

$$(d_{n+1})^r \leq C^*(d_n - d_{n+1}). \tag{31}$$

If $r = 1$, then from (31) we obtain that

$$d_{n+1} \leq \delta d_n \quad \text{with} \quad \delta = \frac{C^*}{1 + C^*}. \tag{32}$$

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

$$d_{n+1} \leq \left(\frac{r-1}{C^*}\xi_0 + d_n^{1-r}\right)^{\frac{1}{1-r}}.$$

Checking on the value of ξ_0 from Lemma 3.2, it is easy to see that

$$c_0 = \frac{(r-1)\xi_0}{C^*} = \frac{r-1}{rd_0^{r-1} + C^*}. \tag{33}$$

By induction, it follows that

$$\begin{aligned} d_{n+1} &\leq \left(2c_0 + d_{n-1}^{1-r}\right)^{-\frac{1}{r-1}} \\ &\leq \cdots \qquad (34) \\ &\leq \left((n+1)c_0 + d_0^{1-r}\right)^{-\frac{1}{r-1}}. \qquad (35) \end{aligned}$$

This proves the theorem for Algorithm 2.1.

Now we proceed with the proof of the theorem for Algorithm 2.2. Notice

$$F(u^n) - F(u^{n+1}) = \sum_{i=1}^{m}\left[F(u^{n+i/m}) - F(u^{n+(i-1)/m})\right]. \tag{36}$$

As $u^{n+\frac{i}{m}}$ is the minimizer of (10), we get by (5)

$$F(u^{n+(i-1)/m}) - F(u^{n+i/m}) \geq \frac{K}{p}\|e_i^n\|_V^p. \tag{37}$$

Thus, estimates (36) and (37) together tell us that

$$F(u^n) \geq F(u^{n+1}) \tag{38}$$

and

$$F(u^n) - F(u^{n+1}) \geq \frac{K}{p} \sum_{i=1}^{m} \|e_i^n\|_V^p. \tag{39}$$

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

$$\begin{aligned} &\langle F'(u^{n+1}) - F'(u), u^{n+1} - u\rangle \\ &= \sum_{i=1}^{m} \left\langle F'(u^{n+1}) - F'(u^{n+(i-1)/m} + e_i^n), v_i \right\rangle \\ &= \sum_{i=1}^{m} \sum_{j\geq i}^{m} \left\langle F'(u^{n+j/m}) - F'(u^{n+(j-1)/m}), v_i \right\rangle \\ &\leq C_2 \left(\sum_{j=1}^{m} \|e_j^n\|_V^p \right)^{\frac{q-1}{p}} \left(\sum_{i=1}^{m} \|v_i\|_V^\sigma \right)^{\frac{1}{\sigma}}. \end{aligned} \tag{40}$$

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

$$\begin{aligned} &\langle F'(u^{n+1}) - F'(u), u^{n+1} - u\rangle \\ &\leq C_2 \left[\left(\sum_{i=1}^{m} \|e_i^n\|_V^p \right)^{\frac{q-1}{p}} \right] \left(\sum_{i=1}^{m} \|v_i\|_V^\sigma \right)^{\frac{1}{\sigma}} \\ &\leq C_1 C_2 \left(\sum_{i=1}^{m} \|e_i^n\|_V^p \right)^{\frac{q-1}{p}} \cdot \|u^{n+1} - u\|_V \\ &\leq C_1 C_2 \left(\frac{p}{K} [F(u^n) - F(u^{n+1})] \right)^{\frac{q-1}{p}} \cdot \|u^{n+1} - u\|_V. \end{aligned} \tag{41}$$

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 $\mathcal{E} = (\epsilon_{ij})$, where ϵ_{ij} satisfies

$$\langle F'(w_{ij} + u_i) - F'(w_{ij}), v_j\rangle \leq \epsilon_{ij} \|u_i\|_V^{q-1} \|v_i\|_V, \forall w_{ij} \in V, \forall u_i \in V_i, \forall v_j \in V_j.$$

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

$$g(t) = F\left(u^n + t \sum_{i=1}^{m} e_i^n \right)$$

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 Ω, and Ω_i has diameter of order H. For each Ω_i, we further divide it into smaller simplices with diameter of order h. In the case that Ω has a curved boundary, we shall also fill the area between $\partial\Omega$ and $\partial\Omega_H$; here $\bar{\Omega}_H = \bigcup_{i=1}^M \bar{\Omega}_i$ with finite elements with diameters of order h. We assume that the resulting elements form a shape regular finite element subdivision of Ω (see Ciarlet [7]). We call this the fine mesh or the h-level subdivision of Ω with mesh parameter h. We denote $\bar{\Omega}_h = \bigcup_{\mathcal{T}\in\mathcal{T}_h} \bar{\mathcal{T}}$ to be the fine mesh subdivision. Let $S_0^H \subset W_0^{1,\infty}(\Omega_H)$ and $S_0^h \subset W_0^{1,\infty}(\Omega_h)$ be the continuous, piecewise rth order polynomial finite element spaces, with zero trace on $\partial\Omega_H$ and $\partial\Omega_h$, over the H-level and h-level subdivisions of Ω, respectively. More specifically,

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

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

For each Ω_i, we consider an enlarged subdomain Ω_i^δ consisting of elements $\mathcal{T}, \mathcal{T} \in \mathcal{T}_h$ with $dist(\mathcal{T}, \Omega_i) \le \delta$. The union of Ω_i^δ covers $\bar{\Omega}_h$ with overlaps of size δ. Let us denote the piecewise rth order polynomial finite element space with zero traces on the boundaries $\partial\Omega_i^\delta$ as $S_0^h(\Omega_i^\delta)$. Then one can show that

$$S_0^h = S_0^H + \sum S_0^h(\Omega_i^\delta). \tag{42}$$

For the overlapping subdomains, assume that there exist m colors such that each subdomain Ω_i^δ 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 $d = 1; m \le 4$ if $d = 2; m \le 8$ if $d = 3$ (see Figure 1). Let Ω_i' be the union of the subdomains with the ith color and

$$V_i = \{v \in S_0^h | \quad v(x) = 0, \quad x \notin \Omega_i'\}.$$

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 $\{\Omega_i'\}_{i=1}^m$, i.e., $\theta_i \in C_0^\infty(\Omega_i' \cap \Omega)$, $\theta_i \ge 0$ and $\sum_{i=1}^m \theta_i = 1$. It can be chosen so that

$$|\nabla\theta_i| \le C/\delta, \quad \theta_i(x) = \begin{cases} 1 & \text{if distance } (x, \partial\Omega_i') \ge \delta \text{ and } x \in \Omega_i', \\ 0 & \text{on } \overline{\Omega\backslash\Omega_i'}. \end{cases}$$

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

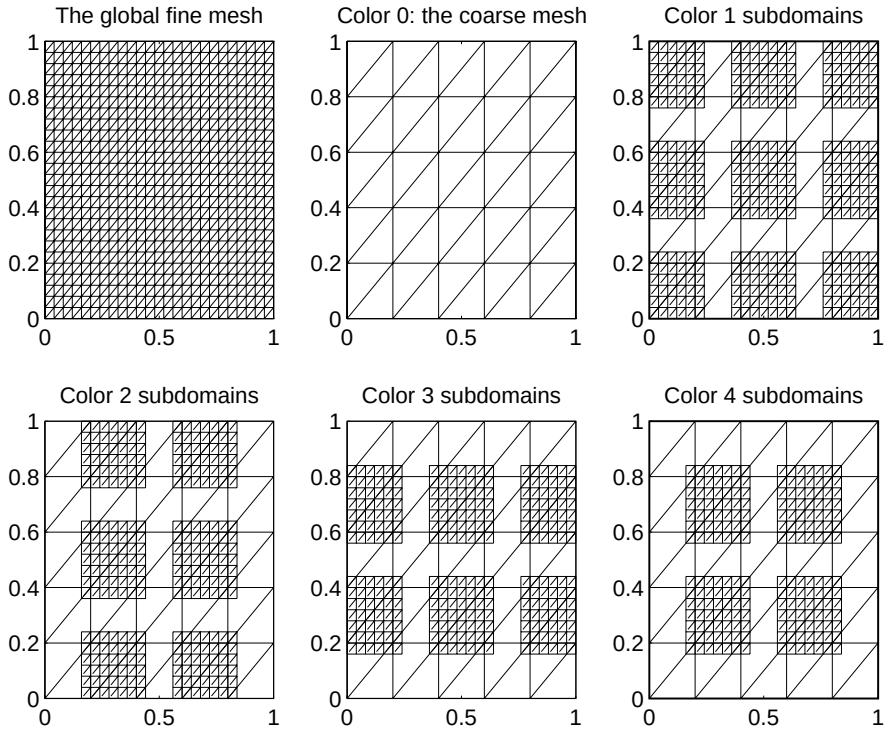

FIGURE 1. The coloring of the subdomains and the coarse mesh grid

Lemma 4.1. *For any $s \geq 1$,*

$$\left(\|v_0\|_{1,p}^s + \sum_{i=1}^m \|v_i\|_{1,p}^s \right)^{\frac{1}{s}} \leq C(m+1)^{\frac{1}{s}} \left(1 + \left(\frac{H}{\delta} \right)^{\frac{p-1}{p}} \right) \|v\|_{1,p}. \tag{44}$$

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:

$$\int_\Omega |\nabla v_0|^p dxdy \leq C \int_\Omega |\nabla v|^p dxdy.$$

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

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

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 $\mathcal{T}$ is constructed by a successive refinement process. More precisely, $\mathcal{T} = \mathcal{T}_J$ for some $J > 1$, and $\mathcal{T}_j$ for $j \le J$ are a nested sequence of quasi-uniform finite element partitions, i.e., $\mathcal{T}_j$ consist of finite elements $\mathcal{T}_j = \{\tau_j^i\}$ of size h_j such that $\Omega = \bigcup_i \tau_j^i$ for which the quasi-uniformity constants are independent of j (cf. [7]), and τ_{j-1}^l 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 γ^{2j}.

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 $\mathcal{T}_1$ being the given coarsest initial triangulation, which is quasi-uniform. In this example, $\gamma = 1/\sqrt{2}$. We can use much smaller γ in constructing the meshes, but the constant C_1 is getting larger when γ is becoming smaller (see (46)).

Corresponding to each finite element partition $\mathcal{T}_j$, a finite element space $\mathcal{M}_j$ can be defined by

$$\mathcal{M}_j = \{v \in W^{1,p}(\Omega) : v|_\tau \in \mathcal{P}_1(\tau), \quad \forall\, \tau \in \mathcal{T}_j\}.$$

Each finite element space $\mathcal{M}_j$ 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 $\mathcal{T}_j$. Associated with each such nodal basis function, we define a one dimensional subspace as

$$\mathcal{M}_j^i = \text{span } (\phi_j^i).$$

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 $V_j^k,\ k = 1, 2, \cdots, m_c$ be the sum of the subspaces $\mathcal{M}_j^i$ associated with nodes of the kth color on level j. Letting $V = \mathcal{M}_J$, we have the following trivial space decomposition:

$$V = \sum_{j=1}^{J} \sum_{k=1}^{m_c} V_j^k. \tag{45}$$

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

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

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

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

$$v_j = \sum_{k=1}^{m_c} v_j^k = \sum_{i=1}^{n_j} \nu_j^i.$$

Denote Ω_j^k the union of the support sets of the basis functions associated with the kth color nodes on level j. We estimate

$$\sum_{k=1}^{m_c} \left|v_j^k\right|_{1,p}^{\sigma} = \sum_{k=1}^{m_c} \left(\sum_{x_j^i \in \Omega_j^k} \left|v_j(x_j^i)\right|^p \left|\phi_j^i\right|_{1,p}^p \right)^{\frac{\sigma}{p}} \leq C h_j^{\frac{\sigma(d-p)}{p}} \sum_{k=1}^{m_c} \left(\sum_{x_j^i \in \Omega_j^k} \left|v_j(x_j^i)\right|^p \right)^{\frac{\sigma}{p}}.$$

In the above, we have assumed that $\Omega \subset R^d, d = 1, 2, 3, \cdots$. 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

$$\sum_{k=1}^{m_c} \left|v_j^k\right|_{1,p}^{\sigma} \leq C h_j^{\frac{\sigma(d-p)}{p}} (m_c)^{1-\frac{\sigma}{p}} \left(\sum_{i=1}^{n_j} \left|v_j(x_j^i)\right|^p \right)^{\frac{\sigma}{p}} \leq C h_j^{-\sigma} \|v_j\|_{0,p}^{\sigma}.$$

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

$$\|v_j\|_{0,p}^p \cong h_j^d \sum_{i=1}^{n_j} \left|v_j(x_j^i)\right|^p.$$

As a consequence,

$$\begin{aligned}
\sum_{j=1}^{J} \sum_{k=1}^{m_c} \|v_j^k\|_{1,p}^{\sigma} &\leq C \sum_{j=1}^{J} h_j^{-\sigma} \|v_j\|_{0,p}^{\sigma} \\
&\leq C \sum_{j=1}^{J} h_j^{-\sigma} \left\| \left(Q_j - Q_{j-1}\right) v \right\|_{0,p}^{\sigma} \leq C \sum_{j=1}^{J} h_j^{-\sigma} \left\| Q_j \left(I - Q_{j-1}\right) v \right\|_{0,p}^{\sigma} \\
&\leq C \sum_{j=1}^{J} h_j^{-\sigma} \left\| \left(I - Q_{j-1}\right) v \right\|_{0,p}^{\sigma} \leq C \sum_{j=1}^{J} h_j^{-\sigma} h_{j-1}^{\sigma} \|v\|_{1,p}^{\sigma} \qquad (46) \\
&\leq C \gamma^{-2\sigma} J \|v\|_{1,p}^{\sigma},
\end{aligned}$$

which proves that

$$C_1 \cong J^{\frac{1}{\sigma}} \cong |\log h|^{\frac{1}{\sigma}}.$$

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

$$\langle F'(w_{ij} + u_i) - F'(w_{ij}), v_j \rangle \leq L \|u_i\|_V^{q-1} \|v_j\|_V. \qquad (47)$$

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

(48)

$$\langle F'(w + u) - F'(w), v \rangle \leq L \|u\|_{1,p,supp(u) \cap supp(v)}^{q-1} \|v\|_{1,p,supp(u) \cap supp(v)}.$$

In the above and also later, *supp* denotes the support set of a function. For any $u \in \mathcal{M}_j^i$ and $v \in \mathcal{M}_l^k$, $j \leq l$, 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,

$$\|u\|_{1,p,supp(u) \cap supp(v)} \leq C \gamma^{\frac{2d}{p}|j-l|} \|u\|_{1,p}, \quad \forall u \in \mathcal{M}_j^i, \forall v \in \mathcal{M}_l^k. \qquad (49)$$

Let $w \in V, u \in V_j^i$ and $v \in V_l^k$. We decompose u and v as

$$u = \sum_{\alpha=1}^{n_j} u_\alpha, \quad u_\alpha = u(x_j^\alpha)\phi_j^\alpha, \quad v = \sum_{\beta=1}^{n_k} v_\beta, \quad v_\beta = v(x_l^\beta)\phi_l^\beta,$$

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:

(50)
$$\langle F'(w + \sum u_\alpha) - F'(w), \sum v_\beta\rangle \le \sum_\alpha \sum_\beta \langle F'(w + u_\alpha) - F'(w), v_\beta\rangle.$$

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), (48), (49) and the orthogonality of the one dimensional subspaces of the same color, it is easy to see that

$$\begin{aligned}\langle F'(w+u) - F'(w), v\rangle &\le C\gamma^{\frac{2d(q-1)}{p}(l-j)} L \sum_\alpha \sum_\beta \|u_\alpha\|_{1,p}^{q-1} \|v_\beta\|_{1,p} \\ &= C\gamma^{\frac{2d(q-1)}{p}(l-j)} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V_j^i,\ v \in V_l^k,\ j \le l.\end{aligned}$$

For $j > l$, we shall have

$$\langle F'(w+u) - F'(w), v\rangle \le C\gamma^{\frac{2d}{p}(j-l)} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V_j^i,\ v \in V_l^k.$$

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

(51)
$$\langle F'(w+u) - F'(w), v\rangle \le C{\gamma_0}^{|j-l|} L \|u\|_{1,p}^{q-1} \|v\|_{1,p}, \quad \forall u \in V_j^i,\ v \in V_l^k.$$

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*

$$\|Ax\|_{\ell^\sigma} \le \max_j \left(\sum_i |\theta_{ij}|^\sigma\right)^{\frac{1}{\sigma}} \max_i \left(\sum_j |A_{ij}|^{\sigma'}\right)^{\frac{1}{\sigma'}} \|x\|_{\ell^\sigma}.$$

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

$$\begin{aligned}\|Ax\|_{\ell^\sigma}^\sigma &= \sum_i \left|\sum_j A_{ij}\theta_{ij}x_j\right|^\sigma \\ &\le \sum_i \left(\sum_j |A_{ij}|^{\sigma'}\right)^{\frac{\sigma}{\sigma'}} \left(\sum_j |\theta_{ij}|^\sigma |x_j|^\sigma\right) \\ &\le \max_i \left(\sum_j |A_{ij}|^{\sigma'}\right)^{\frac{\sigma}{\sigma'}} \sum_j \left(\sum_i |\theta_{ij}|^\sigma |x_j|^\sigma\right) \\ &\le \max_j \left(\sum_i |\theta_{ij}|^\sigma\right) \max_i \left(\sum_j |A_{ij}|^{\sigma'}\right)^{\frac{\sigma}{\sigma'}} \sum_j |x_j|^\sigma,\end{aligned}$$

which proves the lemma. □

As a consequence of Lemma 4.3, we easily get the following corollary which generalizes a well-known result from linear algebra (see [25, p.3-38]).

Corollary 4.1. *Let $A = \{A_{ij}\}$ be a symmetric matrix; then*

$$\|Ax\|_{\ell^\sigma} \le \left(\max_i \sum_j |A_{ij}|\right)\|x\|_{\ell^\sigma}.$$

Proof of Corollary 4.1. It is easy to see that

$$|A_{ij}| = |A_{ij}|^{\frac{1}{\sigma'}}|A_{ij}|^{\frac{1}{\sigma}}.$$

The Corollary is an easy consequence of Lemma 4.2 by setting $A_{ij} := |A_{ij}|^{\frac{1}{\sigma'}}$ and $\theta_{ij} := |A_{ij}|^{\frac{1}{\sigma}}$. □

For a given $u_j^i \in V_j^i$ and $v_l^k \in V_l^k$, an application of (4) gives

$$\sum_{i=1}^{m_c}\sum_{k=1}^{m_c}\|u_j^i\|_{1,p}^{q-1}\|v_l^k\|_{1,p} \le m_c\left(\sum_{i=1}^{m_c}\|u_j^i\|_{1,p}^p\right)^{\frac{q-1}{p}}\left(\sum_{k=1}^{m_c}\|v_l^k\|_{1,p}^\sigma\right)^{\frac{1}{\sigma}}.$$

Using the above inequality and Corollary 4.1, we get that

$$\begin{aligned}
&\sum_{j=1}^{J}\sum_{l=1}^{J}\sum_{i=1}^{m_c}\sum_{k=1}^{m_c}\gamma_0^{|j-l|}\|u_j^i\|_{1,p}^{q-1}\ \|v_l^k\|_{1,p}\\
&\le Cm_c\sum_{j=1}^{J}\sum_{l=1}^{J}\gamma_0^{|j-l|}\left(\sum_{i=1}^{m_c}\|u_j^i\|_{1,p}^p\right)^{\frac{q-1}{p}}\left(\sum_{k=1}^{m_c}\|v_l^k\|_{1,p}^\sigma\right)^{\frac{1}{\sigma}} \qquad (52)\\
&\le Cm_c\left(\max_j\sum_{l=1}^{J}\gamma_0^{|j-l|}\right)\left(\sum_{j=1}^{J}\sum_{i=1}^{m_c}\|u_j^i\|_{1,p}^p\right)^{\frac{q-1}{p}}\cdot\left(\sum_{l=1}^{J}\sum_{k=1}^{m_c}\|v_l^k\|_{1,p}^\sigma\right)^{\frac{1}{\sigma}}\\
&\le \frac{Cm_c}{1-\gamma_0}\left(\sum_{j=1}^{J}\sum_{i=1}^{m_c}\|u_j^i\|_{1,p}^p\right)^{\frac{q-1}{p}}\cdot\left(\sum_{l=1}^{J}\sum_{k=1}^{m_c}\|v_l^k\|_{1,p}^\sigma\right)^{\frac{1}{\sigma}}.
\end{aligned}$$

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 (45).

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

$$V = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\mathcal{M}_j^i, \qquad (53)$$

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 [6, 26] and also some other nonlinear partial differential equations for minimum surface, superconductivity and porous media flows.

We consider the following nonlinear problem:

$$\begin{cases} - \quad \nabla \cdot (|\nabla u|^{s-2}\nabla u) = f \text{ in } \Omega \subset R^d \ (1 < s < \infty), \\ \quad\quad u = 0 \text{ on } \partial\Omega. \end{cases} \tag{54}$$

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

$$\min_{v \in W_0^{1,s}(\Omega)} \left[\frac{1}{s} \int_\Omega |\nabla v|^s - \langle f, v \rangle \right].$$

Even with very smooth data, the solution u may not be in the space $W_0^{2,s}$ (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

$$V = W_0^{1,s}(\Omega), \quad F(v) = \int_\Omega (\frac{1}{s}|\nabla v|^s - fv)dx$$

the following estimates:

$$\langle F'(v) - F'(w), v - w \rangle \geq \|v - w\|_{1,s}^s, \quad \text{if } s \geq 2; \tag{55}$$

$$\langle F'(v) - F'(w), v - w \rangle \geq \alpha \frac{\|v - w\|_{1,s}^2}{(\|v\|_{1,s} + \|w\|_{1,s})^{2-s}}, \quad \text{if } 1 < s \leq 2; \tag{56}$$

$$\|F'(v) - F'(w)\|_{V'} \leq \beta(\|v\|_{1,s} + \|w\|_{1,s})^{s-2}\|v - w\|_{1,s}, \quad \text{if } s \geq 2; \tag{57}$$

$$\|F'(v) - F'(w)\|_{V'} \leq \beta\|v - w\|_{1,s}^{s-1}, \quad \text{if } 1 < s \leq 2. \tag{58}$$

In the above, α and β 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:

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

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.,

$$d_n \leq (Cn)^{-\frac{2}{s(s-1)-2}}, \quad \text{for } s > 2;$$

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

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

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

$$\min_{v \in W_0^{1,s}(\Omega)} \int_\Omega \frac{1}{2} a(|\nabla v|^2) + f(v). \tag{59}$$

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

[1] A. Axelsson and W. Layton. A two-level method for the discretization of nonlinear boundary value problems. *SIAM J. Numer. Anal.*, 33:2359–2374, 1996. MR **98c:**65181

[2] R. Bank and D. J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39:453–465, 1982. MR **83j:**65105

[3] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57:1–21, 1991. MR **92d:**65094

[4] X.-C. Cai and M. Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In D. E. Keyes and Xu, editors, *Proceeding of the seventh international conference on domain decomposition methods in Science and scientific computing (Penn. State Univ.)*, pages 21–28. AMS, Providence, 1994. MR **95j:**65157

[5] J. Cea. *Optimisation – théorie et algorithmes*. Dunod, 1971. MR **45:**7941

[6] T. F. Chan and I. Sharapov. Subspace correction multilevel methods for elliptic eigenvalue problems. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *Domain decomposition methods in science and engineering, 9th international conference, Bergen, Norway.*, pages 311–317. DDM.org, 1998. Available online at `http://www.ddm.org/DD9/`.

[7] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978. MR **58:**25001

[8] P. Deuflhard and M. Weiser. Global inexact Newton multilevel FEM for nonlinear elliptic problems. In W. Hackbusch and G. Wittum, editors, *Multigrid methods V*, Lecture Notes in Computational Science and Engineering, Stuttgart, 1996, Springer. CMP 2000:06

[9] J. Douglas, T. Dupont, and L. Wahlbin. The stability in L^q of the L^2-projection into finite element function spaces. *Numer. Math.*, 23:193–197, 1975. MR **52:**4669

[10] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, 37:296–311, 1997. MR **97m:**65100

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

[12] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North–Holland, Amsterdam, 1976. MR **57:**3931b

[13] E. Gelman and J. Mandel. On multilevel iterative methods for optimization problems. *Math. Programming*, 48:1–17, 1990. MR **91b:**90180

[14] R. Glowinski and A. Marrocco. Sur l'approximation par éléments finis d'ordre un, et la résolution par pénalisation-dualité, d'une classe de problémes de Dirichlet non linéaires. *Rev. Fr. Autom. Inf. Rech. Oper. Ser. Rouge Anal. Numér.* 9: 41–76, 1975. MR **52:**9645

[15] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70:163–180, 1995. MR **96a:**65164

[16] W. Hackbusch and A. Reusken. Analysis of a damped nonlinear multilevel method. *Numer. Math.*, 55:225–246, 1989. MR **90f:**65194

[17] R. Kornhuber. *Adaptive monotone multigrid methods for nonlinear variational problems.* Teubner, Stuttgart, 1997. MR **98e:**65054

[18] R. Kornhuber. Globally convergent multigrid methods for porous medium type equations. *Preprint*, 1999.

[19] P. L. Lions. On the Schwarz alternating method. I. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42, Philadelphia, PA, 1988. SIAM. MR **90a:**65248

[20] J. Mandel. Algebric study of a multigrid method for some free boundary problems. *Comptes Rendus Academic of Science Paris, Series I*, 298:77–95, 1984. MR **86i:**49035

[21] S. F. McCormick. *Multilevel projection methods for partial differential equations*, volume 62 of *CBMS-NSF regional conference series in applied mathematics.* SIAM, Philadelphia, 1992. MR **93b:**65137

[22] S. Oualibouch and N. el Mansouri. Proximal domain decomposition algorithms and applications to elliptic problems. In R. Glowinski, J. Periaux, Z-C. Shi, and O. Widlund, editors, *Domain decomposition methods in sciences and engineering, 8th international conference, Beijing , China*, pages 91–98. Wiley, 1997.

[23] R. Rannacher. On the convergence of the Newton-Raphson method for strongly nonlinear finite element equations. In P. Wriggers and W. Wagner, editors, *Nonlinear computational mechanics*, pages 11–30. Springer-Verlag, 1991.

[24] A. Reusken. Convergence of the multilevel full approximation scheme including the V–cycle. *Numer. Math.*, 53:663–686, 1988. MR **89k:**65072

[25] H. Schneider, editor. *Recent advances in matrix theory.* The University of Wisconsin press, 1964. MR **29:**2265

[26] I. Sharapov. Multilevel subspace correction for large scale optimization problems. Technical Report CAM-97-31, University of California at Los Angeles, 1997.

[27] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain decomposition: Parallel multilevel methods for elliptic partial differential equations.* Cambridge Univ. Press, Cambridge, 1996. MR **98g:**65003

[28] X.-C. Tai. Parallel function decomposition and space decomposition methods with applications to optimisation, splitting and domain decomposition. *Report No. 231–1992, Institut für Mathematik, TechnischeUniversität Graz*, 1992. Available online at `http://www.mi.uib.no/~tai`.

[29] X.-C. Tai. Parallel function and space decomposition methods. In P. Neittaanmäki, editor, *Finite element methods, fifty years of the Courant element, Lecture notes in pure and applied mathematics*, volume 164, pages 421–432. Marcel Dekker inc., 1994. Available online at `http://www.mi.uib.no/~tai`. CMP 95:03

[30] X.-C. Tai. Domain decomposition for linear and nonlinear elliptic problems via function or space decomposition. In D. Keyes and J. Xu, editors, *Domain decomposition methods in scientific and engineering computing (Proc. of the 7th international conference on domain decomposition, Penn. State University, 1993)*, pages 355–360. American Mathematical Society, 1995. MR **95k:**65121

[31] X.-C. Tai. Parallel function and space decomposition methods – Part I. function decomposition. *Beijing Mathematics*, 1, part 2:104–134, 1995. Available online at `http://www.mi.uib.no/~tai`.

[32] X.-C. Tai. Parallel function decomposition and space decomposition methods: Part II. space decomposition. *Beijing Mathematics*, 1, part 2:135–152, 1995. Available online at `http://www.mi.uib.no/~tai`.

[33] X.-C. Tai and M. Espedal. Rate of convergence of some space decomposition method for linear and nonlinear elliptic problems. *SIAM J. Numer. Anal.*, 35:1558–1570, 1998. MR **99k:**65101

[34] X.-C. Tai and J. Xu. Global convergence of subspace correction methods for convex optimization problems. Report no. 114, Department of Mathematics, University of Bergen, Norway, 1998. Available online at `http://www.mi.uib.no/~tai`.

[35] O. B. Widlund. Some Schwarz methods for symmetric and nonsymmetric elliptic problems. In *Proceedings of the fifth international sympsium on domain decomposition methods for partial differential equations, Norfolk, May, 1991*, Philadelphia, 1992. SIAM. MR **93j:**65202

[36] J. Xu. *Theory of Multilevel Methods*. PhD thesis, Cornell University, 1989.

[37] J. Xu. Iteration methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992. MR **93k:**65029

[38] J. Xu. A novel two-grid method for semilinear elliptic equations. *SIAM J. Sci. Comp.*, 15:231–237, 1994. MR **94m:**65178

[39] J. Xu. Two-grid discretization techniques for linear and nonlinear PDEs *SIAM J. Numer. Anal.*, 27:1759–1777, 1996. MR **97i:**65169

[40] J. Zou. A new fast solver – monotone MG method (MMG). *J. Comp. Math.*, 5:325–335, 1987. MR **89k:**65148

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/`