Department
of
APPLIED MATHEMATICS

Rate of Convergence for some constraint Dcomposition methods for nonlinear variational inequalities.

by

Xue-Cheng Tai

Report no. 150 October 2000

UNIVERSITY OF BERGEN
Bergen, Norway

Department of Mathematics
University of Bergen
5008 Bergen
Norway

ISSN 0084-778x

Rate of Convergence for some constraint decompostion methods for nonlinear variational inequalities.

by

Xue-Cheng Tai

Report no. 150

October 2000

Rate of Convergence for some constraint decomposition methods for nonlinear variational inequalities

Xue–Cheng Tai*

October 20, 2000

1 Introduction

In this work, we extend the space decomposition and subspace correction algorithms of [61, 57] to solve convex optimization problems over a convex constraint subset. One of the main concerns of this work is the rate of convergence when multilevel domain decomposition and multigrid methods are used to solve some obstacle problems.

From the time that multigrid and domain decomposition methods have gotten the attention of numerical mathematicians and engineers, efforts have been continuously devoted to the study of using domain decomposition and multigrid methods for obstacle problems, see [2, 4, 3, 11, 19, 25, 30, 28, 29, 31, 32, 26, 27, 36, 34, 35, 37, 38, 39, 43, 42, 22, 44, 47, 49, 59, 52, 53, 51, 50, 54, 56, 62]. In the book of McCormick [44, p.100], treatment of constraints for multilevel methods was listed as one of the open and challenging problems. For linear elliptic partial differential equations, it is known that the solution will be influenced globally if the boundary value or the right hand is perturbed around a point. This justifies the need for coarser meshes when using iterative solvers to solve the problems. However, this is not the case for obstacle type problems. A small perturbation of the input data may only influence a small part of the solution domain due to the appearance of the obstacles. Related to this difficulty, the algorithms in [25, 30, 32] are trying to use the active set strategy to separate the obstacle from the solving of the partial differential equations, i.e. during the iterative procedure, the algorithms are trying to identify the active regions of the obstacles and then solve a partial differential equation where the obstacle is not active. The algorithms proposed in [2, 3, 26, 37, 39, 54, 59, 62] are specified for domain decomposition methods. Due to the absence of the coarse mesh in the algorithms, the convergence of the algorithms depends on the number of subdomains. One of the contributions of this work is the convergence rate estimates. For the obstacle problem, it is shown that the algorithms have a convergence rate which is of the same order as the linear unconstrained elliptic problems.

To be more precise, we classify the main contributions of this work into the following few points:

- Convergence for obstacle problems for overlapping domain decomposition methods without a coarse mesh has been studied in many papers. Rate of convergence has been studied in [4, 62, 54]. However, all of these convergence

*Department of Mathematics, University of Bergen, Johannes Brunsgate 12, 5007, Bergen, Norway. Email: `tai@mi.uib.no` and URL: `http://www.mi.uib.no/~tai`. This work was partially supported by the Norwegian Research Council under projects 128224/431 and SEP-115837/431.

proofs require that the computed solutions increase or decrease monotonically to the true solution. Numerical evidence has shown that linear convergence is correct even if the computed solution is not monotonically increasing or decreasing. In this work, we show that the overlapping domain decomposition method has a linear convergence rate which is of the same order as the unconstrained case if the obstacle and the computed functions are decomposed correctly.

- Numerical experiments and convergence analysis for the two-level domain decomposition method, i.e. an overlapping domain decomposition with a coarse mesh, seem still missing in the literature. The real difficulty is the determination of the coarse mesh obstacle. It shall be shown that the algorithm may not converge or converges as slow as the one-level method if the obstacle and the computed solutions are not decomposed properly. In this work, a linear convergence with a convergence rate independent of the mesh parameters and the number of subdomains is obtained by using a proper decomposition of the obstacle and iterative solutions. The nonlinear interpolation operator $I_H^{\ominus}$ defined in §4 play an important role in the decomposition. Moreover, our algorithms are different from the literature ones.
- Multigrid method has been intensively studied for obstacle problems. Convergence has been studied in [11, 25, 30, 28, 32, 22, 36, 42, 43, 44] without analyzing the rate of convergence. Asymptotic linear convergence rate estimates for multigrid methods can be found in Kornhuber [34, 35] which can be regarded as the pioneering works for multigrid convergence rate analysis. We propose some different algorithms for multigrid method. A linear convergence rate is proved for the proposed algorithms. Moreover, the convergence estimates are valid right from the first iteration. We do not need to assume that the obstacle problem is nondegenerate (c.f. [43, p.84], [34, p.173]) and also do not need to assume that the contact region between the obstacle and the true solution has been identified, see [34, p.173, Lemma 2.2], [35]. The convergence rate is valid for all kind of obstacles from $H^1(\Omega)$.
- In applications to domain decomposition and multigrid methods, the method we use to get the obstacle functions for the subproblems is really different from the methods given in the afore mentioned references. We propose to decompose the global obstacle function. In order to get a mesh independent linear convergence, the initial function must be decomposed properly. In the literature, the global obstacle is often used for the subproblems.

Even though our main concern is the obstacle problem, our algorithms are presented in a general setting for general space decompositions. The general algorithms as well as the assumptions are given in §2. The convergence analysis for the general algorithms under the given assumptions are stated in §3. The convergence rate depends essentially on two constants, i.e. C_1 and C_2, see (7) and (8). In section §5, we show that domain decomposition and multigrid methods can be interpreted as space decomposition techniques and can be used for solving the obstacle problems.

2 The optimization problem and the algorithms

2.1 The optimization problem

Given a reflexive Banach space V and a convex functional $F : V \mapsto R$, we shall consider the following nonlinear optimization problem

$$\min_{v\in K} F(v), \quad K \subset V \,. \tag{1}$$

The nonempty convex subset K is assumed to be closed in the strong topology of V. We are interested in the case that the space V and the convex set K can be decomposed as:

$$V = \sum_{i=1}^{m} V_i, \quad K = \sum_{i=1}^{m} K_i, \quad K_i \subset V_i \ . \tag{2}$$

We assume that the functional F is Gâteaux differentiable (see [12]) and that there exists a constant $\kappa > 0$ such that

$$\langle F'(w) - F'(v), w - v \rangle \geq \kappa \|w - v\|_V^2, \quad \forall w, v \in V \ . \tag{3}$$

Here $\langle \cdot, \cdot \rangle$ is the duality pairing between V and its dual space V', i.e. the value of a linear function at an element of V. Under the assumption (3), problem (1) has a unique solution, see [17, p. 35]. For some nonlinear problems, the constant κ may depend on v and w.

The general theory developed for (1) will be applied to the following obstacle problem in connection with finite element approximations:

$$\text{Find } u \in K, \text{ such that } \ a(u, v - u) \geq f(v - u), \quad \forall v \in K, \tag{4}$$

with

$$a(v, w) = \int_\Omega \nabla v \cdot \nabla w \, dx, \quad K = \{ v \in H_0^1(\Omega) | \ v(x) \geq \psi(x) \text{ a.e. in } \Omega \}. \tag{5}$$

It is well known that the above problem is equivalent to the following minimization problem

$$\min_{v \in K} F(v), \quad F(v) = \frac{1}{2} a(v, v) - f(v), \tag{6}$$

assuming that $f(v)$ is a linear functional on $H_0^1(\Omega)$. For simplicity, the domain $\Omega \subset R^d$ is assumed to be bounded and is a polygonal (d=2) or polyhedral (d=3) domain.

For the obstacle problem (4), the minimization space $V = H_0^1(\Omega)$. Correspondingly, we have $\kappa = 1$ for assumption (3).

For simplicity, $\| \cdot \|$ shall be used for the norm of V. Standard notations for Sobolev spaces $H^k(\Omega)$ and $W^{k,p}(\Omega)$ will be used, i.e. $\| \cdot \|_{k,p,D}$ denotes the $W^{k,p}$-norm on a domain D, and $\| \cdot \|_{k,D}$ denotes the H^k-norm on a domain D. The semi-norms are denoted by $| \cdot |_{k,D}$ and $| \cdot |_{k,p,D}$. In the case $D = \Omega$, we will omit D. The generic positive constant C, which may differ from context to context, will be used to denote a constant that is independent of the variables appearing in the inequalities or equalities and the size of the finite element meshes.

Obstacle problems arise from many important applications. For some concrete examples, we refer to Baiocchi and Capelocite [5], Cottle et al. [14], Duvaut and Lions [16], Elliot and Ockendon [18], Glowinski [23], Glowinski et al. [24], Kinderlehrer and Stampaccia [33], Kornhuber [36], and Rodrigues [46]. See also [1, 21, 48] for some recent researches on general iterative methods for linear complementary problems.

2.2 Conditions for the convergence of the algorithms

We need to impose some conditions on the decomposed subspaces to guarantee that the proposed algorithms have a uniform linear convergence rate. First, we assume that there exits a constant $C_1 > 0$ and this constant is fixed once the decomposition (2) is fixed. With such an $C_1 > 0$, it is assumed that any $u, v \in K$ can be decomposed into a sum of $u_i, v_i \in K_i$ and the decompositions satisfy

$$u = \sum_{i=1}^{m} u_i \ , \quad v = \sum_{i=1}^{m} v_i, \quad \text{and} \quad \left(\sum_{i=1}^{m} \|u_i - v_i\|^2 \right)^{\frac{1}{2}} \le C_1 \|u - v\| \ . \tag{7}$$

For given $u, v \in K$, the decompositions u_i, v_i satisfying (7) may not be unique. In addition to the assumption of the existence of such a constant C_1, we also need to assume that there is a $C_2 > 0$ such that

$$\begin{aligned} &\sum_{i=1}^{m} \sum_{j=1}^{m} |\langle F'(w_{ij} + \hat{v}_i) - F'(w_{ij}), \tilde{v}_j \rangle| \\ &\quad \le C_2 \Big(\sum_{i=1}^{m} \|\hat{v}_i\|^2 \Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{m} \|\tilde{v}_j\|^2 \Big)^{\frac{1}{2}} \ , \\ &\quad \forall\ w_{ij} \in V, \forall \hat{v}_i \in V_i \text{ and } \forall \tilde{v}_j \in V_j \ . \end{aligned} \tag{8}$$

2.3 The algorithms

The following algorithms for general space decompositions can be regarded as a generalization of the Jacobi and Gauss-Seidel methods, see [9, 57, 61]. For algorithm 2, all the subproblems shall be computed sequentially. For algorithm 1, all the subproblems are computed in parallel. In applications to domain decomposition methods for linear elliptic partial differential equations without constraints, Algorithm 1 is in fact the additive Schwarz method and Algorithm 2 is the multiplicative Schwarz method. In applications to multigrid methods for linear elliptic partial differential equations without constraints, Algorithm 1 is essentially similar to the ideas used in the BPX preconditioner [10] and Algorithm 2 reduces to sequential multigrid methods. Algorithm 1 is sometimes called the additive or parallel space decomposition method and Algorithm 2 is sometimes called the multiplicative or successive space decomposition method (c.f. [55]).

For a given approximate solution $u \in K$, we shall find a better solution w using the following two algorithms.

Algorithm 1

1. *Choose a relaxation parameter* $\alpha \in (0, 1/m]$ *and decompose* u *into a sum of* $u_i \in K_i$ *satisfying (7).*

2. *Find* $\hat{w}_i \in K_i$ *in parallel for* $i = 1, 2, \cdots, m$ *such that*

$$F\left(\sum_{j=1, j\neq i}^{m} u_j + \hat{w}_i \right) \le F\left(\sum_{j=1, j\neq i}^{m} u_j + v_i \right) \ , \quad \forall v_i \in K_i \ . \tag{9}$$

3. *Set*

$$w_i = (1-\alpha) u_i + \alpha \hat{w}_i \quad \text{and} \quad w = (1-\alpha) u + \alpha \sum_{i=1}^{m} \hat{w}_i \ . \tag{10}$$

Algorithm 2

1. *Choose a relaxation parameter* $\alpha \in (0, 1]$ *and decompose* u *into a sum of* $u_i \in K_i$ *satisfying (7).*

2. Find $\hat{w}_i \in K_i$ sequentially for $i = 1, 2, \cdots, m$ such that

$$F\left(\sum_{j<i} w_j + \hat{w}_i + \sum_{j>i} u_j\right) \le F\left(\sum_{j<i} w_j + v_i + \sum_{j>i} u_j\right), \quad \forall v_i \in K_i\ . \tag{11}$$

and set

$$w_i = (1-\alpha)u_i + \alpha\hat{w}_i.$$

3. Define

$$w = (1-\alpha)u + \alpha\sum_{i=1}^{m}\hat{w}_i\ . \tag{12}$$

In implementations, it may not be necessary to compute and store the values of $\hat{w}_i, w_i$ and u_i. It is possible to define other auxiliary functions and to compute and store these auxiliary functions could make the implementation simpler. For Algorithm 1, under-relaxation (i.e. $\alpha \le 1$) must be introduced in order to guarantee the convergence. Even for the unconstrained case (i.e. $K = V$), the algorithm can diverge when $\alpha > 1$, see Remark 4.1. of [53, p.146]. For Algorithm 2, over-relaxation (i.e. $\alpha > 1$) may accelerate the convergence, but it is hard to do the analysis. In this work, the convergence of Algorithm 2 is only analyzed for the case that $\alpha \le 1$. An analysis for some problems with $K = V$ and $\alpha > 1$ can be found in Frommer and Renaut [20].

3 Convergence Analysis for the Algorithms

Using similar definitions as in [57], we shall use the following notations in the proofs. u^* will always be used to denote the unique solution of (1), which satisfies [17]

$$\langle F'(u^*), v - u^*\rangle \ge 0, \quad \forall v \in K. \tag{13}$$

In addition, we define

$$e_i = \hat{w}_i - u_i, \quad \hat{w} = \sum_{i=1}^{m}\hat{w}_i. \tag{14}$$

The convergence of Algorithms 1 and 2 is given in the following theorem.

Theorem 1 *Assuming that the space decomposition satisfies (7), (8) and that the functional F satisfies (3). Then for Algorithms 1 and 2, we have*

$$\frac{F(w) - F(u^*)}{F(u) - F(u^*)} \le \left(1 - \frac{\alpha}{(\sqrt{1+C^*} + \sqrt{C^*})^2}\right), \tag{15}$$

with

$$C^* = \left(C_2 + \frac{[C_1C_2]^2}{2\kappa}\right)\frac{2}{\kappa}. \tag{16}$$

Proof. Using the notations of (14) and the fact that F is differentiable and convex, it is known (see Ekeland and Temam [17]) that (9) is equivalent to

$$\langle F'(u + e_i), v_i - \hat{w}_i\rangle \ge 0, \quad \forall v_i \in K_i. \tag{17}$$

Under the assumption of (3), it is known that (See Tai and Epsedal [55, Lemma 3.2])

$$F(v_1) - F(v_2) \ge \langle F'(v_2), v_1 - v_2\rangle + \frac{\kappa}{2}\|v_1 - v_2\|^2, \quad \forall v_1, v_2 \in V\ . \tag{18}$$

Define

$$w^{\frac{i}{m}} = \sum_{j=1, j\neq i}^{m} u_j + \hat{w}_i. \tag{19}$$

From (10), we see that $w^{\frac{i}{m}} = u + e_i$ and

$$w = u + \alpha \sum_{i=1}^{m} (\hat{w}_i - u_i) = (1 - \alpha m) u + \alpha \sum_{i=1}^{m} w^{\frac{i}{m}}. \tag{20}$$

Using (17), (20), the convexity of F and (3), and applying similar techniques as in [55, p.1563], it can be proved that

$$\begin{aligned}
& F(u) - F(w) \\
\geq\ & F(u) - \sum_{i=1}^{m} \alpha F(u + e_i) - (1 - \alpha m) F(u) \qquad (21)\\
=\ & \sum_{i=1}^{m} \alpha \Big(F(u) - F(u + e_i) \Big) \\
\geq\ & -\sum_{i=1}^{m} \alpha \langle F'(u + e_i), e_i \rangle + \frac{\kappa}{2} \sum_{i=1}^{m} \alpha \|e_i\|^2 \\
\geq\ & \frac{\kappa}{2} \sum_{i=1}^{m} \alpha \|e_i\|^2 .
\end{aligned}$$

For notational simplicity, we introduce for a given i

$$\phi_j = \begin{cases} u + \displaystyle\sum_{k=i}^{j+i-1} e_k , & \forall j \in [1, m - i + 1] ; \\ u + \displaystyle\sum_{k=i}^{m} e_k + \sum_{k=1}^{j-m+i-1} e_k , & \forall j \in [m - i + 2, m] . \end{cases} \tag{22}$$

It is clear that ϕ_j depends on i. Moreover, we see that

$$\begin{aligned}
\phi_1 &= u + e_i, \\
\phi_2 &= u + e_i + e_{i+1}, \\
&\ldots \\
\phi_m &= u + \sum_{k=1}^{m} e_k.
\end{aligned}$$

It is easy to see that

$$F'\Big(u + \sum_{j=1}^{m} e_j\Big) - F'(u + e_i) = \sum_{j=2}^{m} \left(F'(\phi_j) - F'(\phi_{j-1})\right) . \tag{23}$$

From assumption (7), there exists $u_i^* \in K_i$ such that

$$u^* = \sum_{i=1}^{m} u_i^*, \qquad \Big(\sum_{i=1}^{m} \|u_i - u_i^*\|^2 \Big)^{\frac{1}{2}} \leq C_1 \|u - u^*\|. \tag{24}$$

We shall use (8), (10), (17), (23) and (24) to estimate

$$\langle F'(\hat{w}), \hat{w} - u^* \rangle = \sum_{i=1}^{m} \Big\langle F'(\hat{w}), \hat{w}_i - u_i^* \Big\rangle$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{m} \left\langle F'(\hat{w}) - F'(u+e_i), \hat{w}_i - u_i^* \right\rangle \qquad \text{(using (17))} \\
&= \sum_{i=1}^{m}\sum_{j=2}^{m} \left\langle F'(\phi_j) - F'(\phi_{j-1}), \hat{w}_i - u_i^* \right\rangle \qquad \text{(using (23))} \\
&\leq C_2 \Big(\sum_{j=1}^{m} \|e_i\|^2\Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{m} \|\hat{w}_i - u_i^*\|^2\Big)^{\frac{1}{2}} \qquad \text{(using (8))} \qquad (25) \\
&\leq C_2 \left(\sum_{i=1}^{m} \|e_i\|^2\right)^{\frac{1}{2}} \left(\left(\sum_{i=1}^{m} \|e_i\|^2\right)^{\frac{1}{2}} + C_1 \|u - u^*\| \right) \text{(using (14) and (24))} \\
&= C_2 \sum_{i=1}^{m} \|e_i\|^2 + C_1C_2 \Big(\sum_{i=1}^{m} \|e_i\|^2\Big)^{\frac{1}{2}} \|u - u^*\|.
\end{aligned}
$$

From (3) and (18), it is easy to see that

$$\frac{\kappa}{2}\|u - u^*\|^2 \leq F(u) - F(u^*), \quad F(\hat{w}) - F(u^*) \leq \langle F'(\hat{w}), \hat{w} - u^* \rangle.$$

Let $\mu \in (0,1)$ be a constant to be determined later. We get from the above inequalities, (21), (25) and the inequality $ab \leq \frac{a^2}{4\mu} + \mu b^2$ that

$$
\begin{aligned}
&F(\hat{w}) - F(u^*) \\
&\leq \frac{2C_2}{\alpha\kappa}[F(u) - F(w)] + C_1C_2\sqrt{\frac{2}{\alpha\kappa}}[F(u) - F(w)]^{\frac{1}{2}}\sqrt{\frac{2}{\kappa}}[F(u) - F(u^*)]^{\frac{1}{2}} \\
&\leq \left(C_2 + \frac{[C_1C_2]^2}{2\kappa\mu}\right)\frac{2}{\alpha\kappa}[F(u) - F(w)] + \mu[F(u) - F(u^*)] \qquad (26) \\
&\leq \left(C_2 + \frac{[C_1C_2]^2}{2\kappa}\right)\frac{2}{\alpha\kappa\mu}[F(u) - F(w)] + \mu[F(u) - F(u^*)].
\end{aligned}
$$

From the definition of C^* in (16), we get from the convexity of F, (10), (21) and (26) that

$$
\begin{aligned}
&F(w) - F(u^*) \leq (1-\alpha)F(u) + \alpha F(\hat{w}) - F(u^*) \\
&= (1-\alpha)(F(u) - F(u^*)) + \alpha(F(\hat{w}) - F(u^*)) \\
&\leq (1-\alpha)(F(u) - F(u^*)) + C^*\mu^{-1}(F(u) - F(w)) + \alpha\mu(F(u) - F(u^*)) \\
&= (1 - \alpha + C^*\mu^{-1} + \alpha\mu)(F(u) - F(u^*)) - C^*\mu^{-1}(F(w) - F(u^*)),
\end{aligned}
$$

and thus

$$\frac{F(w) - F(u^*)}{F(u) - F(u^*)} \leq \frac{1 + C^*\mu^{-1} - \alpha + \alpha\mu}{1 + C^*\mu^{-1}} = 1 - \alpha\frac{\mu(1-\mu)}{\mu + C^*} \quad \forall \mu \in (0,1).$$

For a given C^*, the function $g(\mu) = \frac{\mu(1-\mu)}{\mu + C^*}$ has a unique maximizer in $[0,1]$ and the maximizer is $\mu^* = \sqrt{(C^*)^2 + C^*} - C^* \in (0,1)$. Moreover,

$$g(\mu^*) = \frac{1}{(\sqrt{C^*+1} + \sqrt{C^*})^2}.$$

This proves the theorem for Algorithm 1. We are only interested in the case that $C_1C_2 = O(1)$ or $C_1C_2 \gg 1$. In case that $C_1C_2 = o(1)$, the proof can be refined to show that the convergence rate is also of order $o(1)$, i.e. the convergence rate goes to zero when C_1C_2 goes to zero.

To prove the convergence rate for Algorithm 2, define

$$w^{\frac{i}{m}} = \sum_{j\leq i} w_j + \sum_{j>i} u_j, \quad \hat{w}^{\frac{i}{m}} = \sum_{j<i} w_j + \hat{w}_i + \sum_{j>i} u_j. \tag{27}$$

We see that

$$w^0 = u, \quad w^{\frac{m}{m}} = w, \quad w^{\frac{i}{m}} = (1-\alpha)w^{\frac{i-1}{m}} + \alpha\hat{w}^{\frac{i}{m}}, \tag{28}$$

$$F(u) - F(w) = \sum_{i=1}^{m} \left[F(w^{(i-1)/m}) - F(w^{i/m})\right]. \tag{29}$$

Since $\hat{w}^{\frac{i}{m}}$ is the minimizer of (11), it satisfies

$$\langle F'(\hat{w}^{\frac{i}{m}}), v_i - \hat{w}_i\rangle \geq 0, \quad \forall v_i \in K_i. \tag{30}$$

Using (18), (28), (30) and the convexity of F to get

$$F(w^{(i-1)/m}) - F(w^{i/m}) \geq \alpha\left(F(w^{(i-1)/m}) - F(\hat{w}^{i/m})\right) \geq \frac{\alpha\kappa}{2}\|e_i\|^2. \tag{31}$$

Thus, estimates (29) and (31) together lead to

$$F(u) - F(w) \geq \frac{\kappa}{2}\sum_{i=1}^{m}\alpha\|e_i\|^2 \quad \text{and so} \quad F(u) \geq F(w).$$

Similar as in (22), we can introduce functions ϕ_j to satisfy

$$\begin{aligned}
&\phi_j - \phi_{j-1} = \hat{w}_i - w_i = (1-\alpha)e_i, \quad j < i;\\
&\phi_j - \phi_{j-1} = \hat{w}_i - u_i = e_i, \quad j > i; \quad \phi_i - \phi_{i-1} = 0;\\
&F'(\hat{w}) - F'(\hat{w}^{\frac{i}{m}}) = \sum_{j=2}^{m}(F'(\phi_j) - F'(\phi_{j-1})).
\end{aligned}$$

We use (7), (8), (24), (30), the above equalities and the fact that $0 \leq 1-\alpha < 1$ to get

$$\begin{aligned}
&\langle F'(\hat{w}), \hat{w} - u^*\rangle\\
\leq\ &\sum_{i=1}^{m}\left\langle F'(\hat{w}) - F'(\hat{w}^{i/m}), \hat{w}_i - u_i^*\right\rangle\\
=\ &\sum_{i=1}^{m}\sum_{j=2}^{m}\langle F'(\phi_j) - F'(\phi_{j-1}), \hat{w}_i - u_i^*\rangle\\
\leq\ &C_2\left(\sum_{j=1}^{m}\|e_j\|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{m}\|\hat{w}_i - u_i^*\|^2\right)^{\frac{1}{2}}
\end{aligned} \tag{32}$$

The rest of the proof is the same as for Algorithm 1. ■

4 Finite element spaces and some constrained interpolation operators

In this section, we shall propose some interpolation operators subject to some constraints. These operators are not only needed in our analysis for the algorithms, but also needed in the implementation of the algorithms. We use essentially these operators to decompose the constraint sets and functions satisfying assumption (7).

Let $\mathcal{T}_h$ be a quasi-uniform triangulation of the domain Ω with a mesh size h and $S_h \subset H_0^1(\Omega)$ be the corresponding piecewise linear finite element space on $\mathcal{T}_h$ [13]. In the analysis, we need to use finite element spaces with different mesh sizes. It will be assumed that h is always the smallest mesh size. For an $H > h$, we consider the case that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. Operator $I_H : C(\bar{\Omega}) \mapsto S_H$ will always be used to denote the nodal Lagrangian interpolation into S_H for any $H \geq h$.

In the following, the definition of a nonlinear interpolation operator $I_H^{\ominus} : S_h \mapsto S_H$ will be given. Denote by $\mathcal{N}_H = \{x_0^i\}_{i=1}^{no}$ all the interior nodes for $\mathcal{T}_H$. For a given x_0^i, let ω_i be the union of the mesh elements of $\mathcal{T}_H$ having x_0^i as one of its vertices, i.e.

$$\omega_i := \cup\{\tau \in \mathcal{T}_H, x_0^i \in \bar{\tau}\}.$$

Let $\{\phi_0^i\}_{i=1}^{no}$ be the associated nodal basis functions satisfying $\phi_0^i(x_0^k) = \delta_{ik}$, $\phi_0^i \geq 0$, $\forall i$ and $\sum_i \phi_0^i(x) = 1$. It is clear that ω_i is the support of ϕ_0^i. Given a nodal point $x_0^i \in \mathcal{N}_H$ and a $v \in S_h$, let

$$I_i v = \min_{\omega_i} v(x) \tag{33}$$

The interpolated function is then defined as

$$I_H^{\ominus} v := \sum_{x_0^i \in \mathcal{N}_H} (I_i v)\phi_0^i(x).$$

This nonlinear operator is essentially an extension of the intergrid operator used in [34], [42]. From the definition, it is easy to see that

$$I_H^{\ominus} v \leq v, \quad \forall v \in S_h, \tag{34}$$

$$I_H^{\ominus} v \geq 0, \quad \forall v \geq 0, v \in S_h. \tag{35}$$

Moreover, the interpolation for a given $v \in S_h$ on a finer mesh is always bigger than the corresponding interpolation on a coarser mesh due to the fact that each coarser mesh element contains several finer mesh elements, i.e.

$$I_{h_1}^{\ominus} v \leq I_{h_2}^{\ominus} v, \quad \forall h_1 \geq h_2 \geq h, \quad \forall v \in S_h. \tag{36}$$

In addition, the interpolation operator also has the following approximation properties.

Theorem 2 *For any $u, v \in S_h$, it is true that*

$$\|I_H^{\ominus} u - I_H^{\ominus} v - (u - v)\|_0 \leq CH|u - v|_1, \tag{37}$$

$$\|I_H^{\ominus} v \; - \; v\|_0 \leq CH|v|_1, \qquad \|I_H^{\ominus} v\|_1 \leq C\|v\|_1. \tag{38}$$

Proof. As the interpolation operator is nonlinear with respect to v, the well-known Bramble-Hilbert lemma can not be used in this context. It is necessary to use the older Taylor expansion techniques [8, 7]. To prove (37), define on each ω_i

$$\tilde{u} = u - I_i u, \quad \tilde{v} = v - I_i v.$$

From the definition of I_i of (33), it is true that

$$\tilde{u} \geq 0, \ \tilde{v} \geq 0 \text{ in } \omega_i, \quad \text{and} \quad \min_{\bar{\omega}_i} \tilde{u} = 0, \ \min_{\bar{\omega}_i} \tilde{v} = 0. \tag{39}$$

Thus, there must exist a $\xi_0^i \in \bar{\omega}_i$ such that

$$\tilde{u}(\xi_0^i) = \tilde{v}(\xi_0^i). \tag{40}$$

Otherwise, $\tilde{u} > \tilde{v}$ or $\tilde{u} < \tilde{v}$ and it is impossible to have the last two equalities of (39) to hold simultaneously. As $\tilde{u}$ and $\tilde{v}$ are piecewise linear, we get from (40) that

$$\|\tilde{u}-\tilde{v}\|_{0,\omega_i} \leq CH|\tilde{u}-\tilde{v}|_{1,\omega_i} = CH|u-v|_{1,\omega_i}.$$

For a $\tau \in \mathcal{T}_H$ and $\tau \subset \omega_i$, let $|\tau|$ be the measure of τ and

$$a_\tau = \frac{1}{|\tau|}\int_\tau (u-v)dx.$$

The Poincare-Friedrichs inequality gives

$$\|u-v-a_\tau\|_{0,\tau} \leq CH|u-v|_{1,\tau}.$$

From the definition of $I_H^\ominus$, it is true that

$$I_H^\ominus u - I_H^\ominus v = \sum_{x_0^i \in \mathcal{N}_H} (I_i u - I_i v)\phi_0^i(x).$$

On τ, there are only three nonzero terms in the above summation. As $\sum_i \phi_0^i(x) = 1$, the following estimate is correct.

$$\|I_H^\ominus u - I_H^\ominus v - a_\tau\|_{0,\tau} \leq \sum_i \|I_i u - I_i v - a_\tau\|_{0,\tau} \tag{41}$$

$$\leq \sum_i \|I_i u - I_i v - (u-v)\|_{0,\tau} + 3\|u-v-a_\tau\|_{0,\tau} \tag{42}$$

$$\leq CH|u-v|_{1,\omega_i}. \tag{43}$$

As a consequence

$$\|I_H^\ominus u - I_H^\ominus v - (u-v)\|_{0,\tau} \leq \|I_h^\ominus u - I_H^\ominus v - a_\tau\|_{0,\tau} + \|a_\tau - (u-v)\|_{0,\tau} \leq CH|u-v|_{1,\omega_i},$$

and estimate (37) follows. To prove (38), we just need to set $u = 0$ in (37) and use the inverse inequality for functions from S_H. ■

Based on the operator $I_H^\ominus$, it can be seen that the operator $I_H^{2\ominus} : S_h \mapsto S_H$ defined by

$$I_H^{2\ominus} = I_H^\ominus v + I_H^\ominus(v - I_H^\ominus v), \quad \forall v \in S_h,$$

also satisfies the properties (34), (35), (37) and (38). Moreover, it is true that

$$I_H^\ominus v \leq I_H^{2\ominus} v \leq v, \quad \text{and thus} \quad \|I_H^{2\ominus} v - v\|_0 \leq \|I_H^\ominus v - v\|_0, \quad \forall v \in S_h,$$

i.e. $I_H^{2\ominus} v$ approximates v better than $I_H^\ominus v$. Inductively, we can define

$$I_H^{3\ominus} = I_H^{2\ominus} v + I_H^\ominus(v - I_H^{2\ominus} v), \quad I_H^{4\ominus} v = \cdots \quad \forall v \in S_h$$

and each operator $I_H^{k\ominus}$ satisfies the properties (34), (35), (37) and (38).

From theorem 2, it is easy to see that the following is correct.

Theorem 3 *There exists an interpolation operator $I_H^\oplus : S_h \mapsto S_H$ such that*

$$I_H^\oplus v \geq v, \quad \forall v \in S_h,$$
$$I_H^\oplus v \leq 0, \forall v \leq 0, v \in S_h,$$
$$\|I_H^\oplus u - I_H^\oplus v - (u-v)\|_0 \leq CH|u-v|_1,$$
$$\|I_H^\oplus v \; - \; v\|_0 \leq CH|v|_1, \qquad \|I_H^\oplus v\|_1 \leq C\|v\|_1, \forall v \in S_h.$$

Proof. Replace the definition of $I_i v$ given in (33) by the following

$$I_i v = \max_{\bar{\omega}_i} v(x).$$

The rest of the proof uses the same argument as in Theorem 2. ■

5 Space decompositions for $H_0^1(\Omega)$ and K

In this subsection, we show that the overlapping domain decomposition methods and the multigrid methods can be used to decompose a finite element space and the constraint set K for the obstacle problem (4).

5.1 Overlapping domain decomposition methods

Let $\mathcal{T}_H = \{\Omega_i\}_{i=1}^M$ be a quasi-uniform finite element division, or a coarse mesh, of Ω where Ω_i has diameter of order H. We further divide each Ω_i into smaller simplices with diameter of order h. We assume that the resulting finite element partition $\mathcal{T}_h$ form a shape regular finite element subdivision of Ω, see Ciarlet [13]. We call this the fine mesh or the h-level subdivision of Ω with the mesh parameter h. We denote $S_H \subset W_0^{1,\infty}(\Omega)$ and $S_h \subset W_0^{1,\infty}(\Omega)$ be the continuous, piecewise linear finite element spaces over the H-level and h-level subdivisions of Ω respectively. More specifically,

$$S_H = \left\{ v \in W_0^{1,\infty}(\Omega) \middle| \quad v|_{\Omega_i} \in P_1(\Omega_i), \forall i \right\},$$

$$S_h = \left\{ v \in W_0^{1,\infty}(\Omega) \middle| \quad v|_{\tau} \in P_1(\tau), \forall \tau \in \mathcal{T}_h \right\}.$$

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

$$S_h = \sum_{i=1}^M S_h(\Omega_i^\delta) \quad \text{and} \quad S_h = S_H + \sum_{i=1}^M S_h(\Omega_i^\delta) \, . \tag{44}$$

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 $d = 3$. Let Ω_i' be the union of the subdomains with the i^{th} color, and

$$V_i = \{ v \in S_h | \quad v(x) = 0, \quad x \notin \Omega_i' \}, \; i = 1, 2, \cdots, m.$$

By denoting subspaces $V_0 = S_H$, $V = S_h$, we find that decomposition (44) means

$$a). \quad V = \sum_{i=1}^m V_i \qquad \text{and} \qquad b). \quad V = V_0 + \sum_{i=1}^m V_i. \tag{45}$$

Note that the summation index is now from 0 to m instead of from 1 to m when the coarse mesh is added.

For the constraint set K, we shall first decompose ψ as

$$\psi = \sum_{i=1}^m \psi_i, \quad \text{or} \quad \psi = \psi_0 + \sum_{i=1}^m \psi_i, \; \psi_0 \in V_0, \; \psi_i \in V_i. \tag{46}$$

and then define

$$K_0 = \{ v \in V_0 | \; v \ge \psi_0 \}, \quad \text{and} \quad K_i = \{ v \in V_i | \; v \ge \psi_i \}, \; i = 1, 2, \cdots, m. \tag{47}$$

Under condition (46), it is easy to see that (2) is correct. When the coarse mesh is added, the summation index is from 0 to m.

Following [15, 60], let $\{\theta_i\}_{i=1}^m$ be a partition of unity with respect to $\{\Omega_i'\}_{i=1}^m$, i.e. $\theta_i \in V_i$, $\theta_i \geq 0$ and $\sum_{i=1}^m \theta_i = 1$. It can be chosen so that

$$|\nabla\theta_i| \leq C/\delta, \quad \theta_i(x) = \begin{cases} 1 & \text{if } x \in \tau, \ \text{distance } (\tau, \partial\Omega_i') \geq \delta \text{ and } \tau \subset \Omega_i', \\ 0 & \text{on } \overline{\Omega\backslash\Omega_i'}. \end{cases} \tag{48}$$

The partitions θ_i are needed in our implementations to decompose the constraint set K and the value of u.

5.2 Decompositions without the coarse mesh

If we use the overlapping domain decomposition without the coarse mesh, i.e. we use decomposition (45.a), then we will get some domain decomposition algorithms which is essentially the block-relaxation method. Even in the case that $V = R^n$, the analysis for the convergence rate for general convex functional $F : R^n \mapsto R$ and general convex set $K \subset R^n$ is not an trivial matter, see [40, 41] for a survey. In case that the convex constraint set K has more structure, there are more available convergence rate estimate. For the domain decomposition method without the coarse mesh, convergence proof can be found in [49, 51, 53, 37, 59], etc. Linear convergence rate has been proved in [62, 4, 3, 54]. However, all the proofs require that the computed solutions converge to the true solution monotonically. Numerical evidence shows that linear convergence is true even if the computed solutions are not monotonically increasing or decreasing. In the following, we shall use our theory to prove this fact, i.e. we will get a linear convergence rate without requiring the monotonicity of the computed solutions.

For any given $u, v \in S_h$, we decompose u, v and ψ as

$$u = \sum_{i=1}^m u_i, \quad v = \sum_{i=1}^m v_i, \quad \psi = \sum_{i=1}^m \psi_i,$$
$$u_i = I_h(\theta_i u), \quad v_i = I_h(\theta_i v), \quad \psi_i = I_h(\theta_i \psi).$$

In case that $u, v \geq \psi$, it is true that $u_i, v_i \geq \psi_i$. In addition,

$$\sum_{i=1}^m \|u_i - v_i\|_1^2 \leq C\left(1 + \frac{1}{\delta^2}\right)\|u - v\|_1^2,$$

which shows that

$$C_1 \leq C(1 + \delta^{-1}).$$

The decomposition for u and ψ are needed in the implementation. The decomposition for v is only needed for the analysis. It is known that $C_2 \leq m$ with m being the number of colors. From Theorem 1, the following rate is obtained without requiring that the computed solutions increase or decrease monotonically:

$$\frac{F(w) - F(u^*)}{F(u) - F(u^*)} \leq 1 - \frac{\alpha}{1 + C(1 + \delta^{-2})}.$$

For algorithm 2, we can take $\alpha = 1$.

5.3 Decompositions with the two-level method

Numerical experiments and convergence analysis for the two-level domain decomposition method, i.e. an overlapping domain decomposition with a coarse mesh, seem still missing in the literature. The work of [59] is in fact a two-level algebraic approach and the coarse mesh space V_0 is in fact not used. In §6.2, it will be shown that the algorithms may not converge or converges as slow as the one-level method if the coarse mesh obstacle is not given properly. Decomposing the obstacle and the function u properly, a linear convergence rate which is independent of the mesh sizes and the number of subdomain is obtained for the proposed algorithms.

For the obstacle function ψ, there exist $\psi_0 \in V_0$ and $\psi_i \in V_i$, $i = 1, 2, \cdots, m$ such that $\psi = \psi_0 + \sum_{i=1}^m \psi_i$. The decomposition may not be unique. We just pick any of the decompositions. The analysis and the numerical tests show that this does not affect the convergence rate.

For any given $u \in K$, the decomposition for u shall be obtained from the decomposition of ψ and a decomposition of $u - \psi$ as in the following

$$u - \psi = \sigma_0 + \sum_{i=1}^m \sigma_i, \quad \sigma_0 = I_H^{\ominus}(u - \psi), \ \sigma_i = I_h(\theta_i(u - \psi - \sigma_0)). \tag{49}$$

From (34) , (35) and the fact that $u \geq \psi$, it is true that

$$0 \leq \sigma_0 \leq u - \psi \quad \text{and so} \quad \sigma_i \geq 0, \ i = 1, 2, \cdots, m. \tag{50}$$

Combining (49) and the decomposition for ψ, one gets the following decomposition for u

$$u = u_0 + \sum_{i=1}^m u_i \quad u_0 = \psi_0 + \sigma_0, \quad u_i = \psi_i + \sigma_i. \tag{51}$$

As a consequence of (50), it is correct that $u_0 \in K_0$ and $u_i \in K_i$, $i = 1, 2, \cdots, m$. The decompositions for u and ψ are needed for the implementation of the algorithms. For the analysis, we also decompose any $v \in K$ as

$$v = v_0 + \sum_{i=1}^m v_i \quad v_0 = \psi_0 + I_H^{\ominus}(v - \psi), \quad v_i = \psi_i + I_h(\theta_i(v - \psi - I_H^{\ominus}(v - \psi))). \tag{52}$$

It is clear that $v_0 \in K_0$ and $v_i \in K_i$ for any $v \in K$. It follows from Theorem 2 that

$$\|u_0 - v_0\|_1 \leq C\|u - v\|_1. \tag{53}$$

Note that

$$u_i - v_i = I_h\big(\theta_i(u - v - I_H^{\ominus}(u - \psi) + I_h^{\ominus}(v - \psi))\big).$$

Using estimate (37) and similar to the proofs for the unconstrained cases, c.f. [60], and [57], it can be proven that

$$\|u_i - v_i\|_1^2 \leq C\left(1 + \frac{H}{\delta}\right)\|u - v\|_1^2. \tag{54}$$

Thus

$$\left(\|u_0 - v_0\|_1^2 + \sum_{i=1}^m \|u_i - v_i\|_1^2\right)^{\frac{1}{2}} \leq C(m)\left(1 + \left(\frac{H}{\delta}\right)^{\frac{1}{2}}\right)\|u - v\|_1.$$

The estimate for C_2 is known, c.f. [57]. Thus, for the two-level domain decomposition method, we have

$$C_1 = C(m)\left(1 + \frac{\sqrt{H}}{\sqrt{\delta}}\right), \quad C_2 = C(m),$$

where $C(m)$ is a constant only depending on m, but not on the mesh parameters and the number of subdomains. An application of Theorem 1 will show that the following convergence rate estimate is correct:

$$\frac{F(w) - F(u^*)}{F(u) - F(u^*)} \leq 1 - \frac{\alpha}{1 + C(1 + H^{\frac{1}{2}}\delta^{-\frac{1}{2}})}.$$

For algorithm 2, we can choose $\alpha = 1$. When $H \to 0$ and $h \to 0$, we will get a mesh independent linear convergence if the overlapping size δ is chosen to satisfy $H/\delta =$ constant.

5.4 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 $\mathcal{T}_h$ is constructed by a successive refinement process. More precisely, $\mathcal{T}_h = \mathcal{T}_{h_J}$ for some $J > 1$, and $\mathcal{T}_{h_j}$ for $j \le J$ is a nested sequence of quasi-uniform finite element partitions, i.e. $\mathcal{T}_{h_j}$ consist of finite elements $\mathcal{T}_{h_j} = \{\tau_j^i\}$ of size h_j such that $\Omega = \cup_i \tau_j^i$ for which the quasi-uniformity constants are independent of j (cf. [13]) 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 $\mathcal{T}_{h_1}$ 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 (57).

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

$$\mathcal{M}_j = \{v \in W_0^{1,\infty}(\Omega) : v|_\tau \in \mathcal{P}_1(\tau), \quad \forall\, \tau \in \mathcal{T}_{h_j}\}.$$

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 the interior nodes of $\mathcal{T}_j$. Associated with each nodal basis function, we define a one dimensional subspace as follows

$$V_j^i = \text{span }(\phi_j^i).$$

Letting $V = \mathcal{M}_J$, we have the following trivial space decomposition:

$$V = \sum_{j=1}^{J} \sum_{i=1}^{n_j} V_j^i. \tag{55}$$

Each subspace V_j^i is a one dimensional subspace.

For the obstacle function ψ, assume that $\psi_j^i \in V_j^i$ satisfies $\psi = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \psi_j^i$. The choice of the decomposition is not unique. For simplicity for the presentation of the decompositions of u and v, it shall be assumed that

$$\psi = 0, \quad \psi_j^i = 0, \ \forall i, j. \tag{56}$$

In case that the obstacle functions are not zero, one just need to add ψ_j^i to the decompositions of $u - \psi$ and $v - \psi$ to get the decompositions for u and v, see (58) and (59).

For any $v \ge 0$ and $j \le J - 1$, define $v_j = I_{h_j}^{\ominus} v - I_{h_{j-1}}^{\ominus} v \in \mathcal{M}_j$. Let $v_J = v - I_{h_{J-1}}^{\ominus} v \in \mathcal{M}_J$. A further decomposition of v_j is given by

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

It is easy to see that

$$v = \sum_{j=1}^{J} v_j = \sum_{j=1}^{J}\sum_{i=1}^{n_j} v_j^i.$$

For any $u \ge 0$, it shall be decomposed in the same way, i.e.

$$u = \sum_{j=1}^{J}\sum_{i=1}^{n_j} u_j^i, \quad u_j^i = u_j(x_j^i)\phi_j^i, \quad u_j = I_{h_j}^{\ominus} u - I_{h_{j-1}}^{\ominus} u,\ j < J; \quad u_J = u - I_{h_{J-1}}^{\ominus} u.$$

It follows from (34), (35) and (36) that $u_j^i, v_j^i \ge 0$ for all $u, v \ge 0$, i.e.

$$u_j^i, v_j^i \in K_j^i = \{v \in V_j^i : v \ge \psi_j^i\} \quad \text{under condition (56).}$$

We estimate

$$\sum_{i=1}^{n_j} \left|u_j^i - v_j^i\right|_1^2 = \sum_{i=1}^{n_j} \left|u_j(x_j^i) - v_j(x_j^i)\right|^2 \left|\phi_j^i\right|_1^2$$

$$\le C h_j^{d-2} \sum_{i=1}^{n_j} \left|u_j(x_j^i) - v_j(x_j^i)\right|^2 \le C h_j^{-2} \left|u_j - v_j\right|_0^2.$$

Here, we have used the fact that, in the finite element space, an L^2 norm is equivalent to some discrete L^2 norm, namely $\|v_j\|_0^2 \cong h_j^d \sum_{i=1}^{n_j} \left|v_j(x_j^i)\right|^2$. From the definitions of u_j and v_j and estimate (37), it is easy to see that

$$\|u_j - v_j\|_0 \le C(h_j + h_{j-1})|u - v|_1.$$

As a consequence,

$$\sum_{j=1}^{J}\sum_{i=1}^{n_j} \|u_j^i - v_j^i\|_1^2 \le C \sum_{j=1}^{J} h_j^{-2} \|u_j - v_j\|_0^2$$

$$\le C \sum_{j=1}^{J} h_j^{-2} h_{j-1}^2 \left|u - v\right|_1^2 \le C\gamma^{-2} J \left|u - v\right|_1^2, \tag{57}$$

which proves that

$$C_1 \cong \gamma^{-1} J^{\frac{1}{2}} \cong \gamma^{-1} |\log h|^{\frac{1}{2}}.$$

The estimation for C_2 is known, i.e. $C_2 = C(1 - \gamma^d)^{-1}$, see Tai and Xu [57]. Thus for the multigrid method, the error reduction factor for the algorithms is

$$\frac{F(w) - F(u^*)}{F(u) - F(u^*)} \le 1 - \frac{\alpha}{1 + C\gamma^{-2} J}.$$

For unconstrained linear problems, the dependence on J can be removed with much more complicated analysis [45, 6].

In case that the obstacle functions are not zero, one needs to first decompose $u - \psi$ as

$$u - \psi = \sum_{j=1}^{J}\sum_{i=1}^{n_j} \sigma_j^i, \quad \sigma_j^i(x) = \sigma_j(x_j^i)\phi_j^i(x), \tag{58}$$

$$\sigma_j = I_{h_j}^{\ominus}(u - \psi) - I_{h_{j-1}}^{\ominus}(u - \psi),\ j < J; \quad \sigma_J = (u - \psi) - I_{h_{J-1}}^{\ominus}(u - \psi).$$

The decomposition for u, which is needed in the implementation, is then given by

$$u = \sum_{j=1}^{J}\sum_{i=1}^{n_j} u_j^i, \quad u_j^i = \psi_j^i + \sigma_j^i. \tag{59}$$

The decomposition (55) only represents a "half-V-cycle" (or called "\-cycle") multigrid method. In order to produce the full "V-cycle" or "W-cycle" multigrid iteration, we just need to repeat some of the one dimensional subspaces once more or several times more in the decomposition (55). The estimates for C_1 and C_2 can be done in a very similar way.

In decomposition (55), the total number m of subspaces is $m = \sum_{j=1}^{J} n_j$. 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 $\tilde{V}_j^k$, $k = 1, 2, \cdots m_c$ be the sum of the subspaces V_j^i associated with nodes of the k^{th} color on level j. We have the following trivial space decomposition: $V = \sum_{j=1}^{J} \sum_{k=1}^{m_c} \tilde{V}_j^k$. The total number of subspaces for such a decomposition is $m_c J$. Such a decomposition is only needed theoretically. The algorithm produced by this decomposition with Algorithm 1 is the same as the one produced by decomposition (55). For algorithm 2, the resulting schemes for the two-decompositions are different. However, both have a convergence rate independent of the number of subspaces.

6 Implementation issues and some numerical experiments

We shall test our algorithms for the obstacle problem (4) with $\Omega = [-2, 2] \times [-2, 2]$, $f = 0$ and

$$\psi(x, y) = \sqrt{x^2 + y^2} \quad x^2 + y^2 \leq 1, \qquad \psi(x, y) = -1 \quad \text{elsewhere.}$$

With consistent Dirichlet boundary condition, the problem has an analytical solution

$$u^*(x, y) = \begin{cases} \sqrt{1 - x^2 - y^2} & r \leq r^* \\ -(r^*)^2 \ln(r/R)/\sqrt{1 - (r^*)^2} & r \geq r^* \end{cases}$$

where $r = \sqrt{x^2 + y^2}, R = 2$ and $r^* = 0.6979651482...$, which satisfies

$$(r^*)^2 (1 - \ln(r^*/R)) = 1.$$

The subdomain problems are solved by the augmented Lagrangian approach of Tai [58, p.235] with or without the dimensional splitting. Let matrix A be the matrix associated with the bilinear form $a(\cdot, \cdot)$ for the finite element space and b the load vector associated with the linear functional $l(\cdot)$, then u^* and ψ, which now represent the vectors that contain the nodal values of the finite element functions, satisfy (4) if and only if they satisfy (see [14])

$$Au^* \geq b, \quad u^* \geq \psi, \quad (Au^* - b)(u^* - \psi) = 0.$$

The stopping criteria for the subproblems is

$$\| \min(0, Au - b) \|_{\ell^2} + \| \min(0, u - \psi) \|_{\ell^2} + \| (Au - b)(u - \psi) \|_{\ell^2} \leq TOL. \tag{60}$$

The same stooping criteria is used for Algorithms 1 and 2 for the global problem.

Algorithms 1 and 2 are used as iterative solvers, i.e. we take an initial guess and use Algorithms 1 and 2 to get a better approximation and use this newly computed function as the initial guess to compute another better solution and continue in this way. In the plots, en is the H^1-error between the computed solution at the nth iteration and the true finite element solution, see Figure 1. $e0$ is the initial error. In the implementation for the decompositions of §5.2 and §5.3, we need to construct the functions θ_i which are not unique. We have used several choices that satisfy (48) and it seems that they do not alter the convergence rate much.

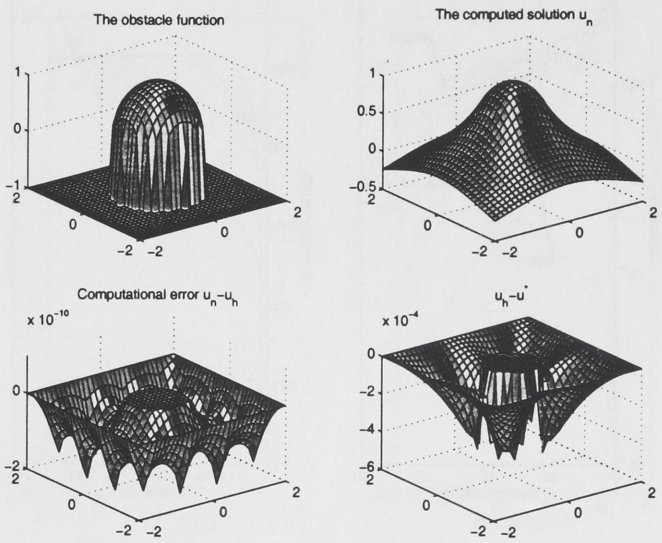

Figure 1: The obstacle and the true finite element solution with $h = 4/128$. u_n is the computed solution by algorithms 1 or 2, u_h is the true finite element solution and u^* is the analytical solution.

6.1 Experiments without the coarse mesh

Without the coarse mesh, the computed solutions will increase monotonically to the true solution if we start with a function which is less than the true solution [54]. We shall start with a function that is less than the true solution in part of the domain and bigger than the true solution in the rest of the domain. Thus, the convergence will not be monotonically. Linear convergence is observed, see Figures 2 and 3. In Figure 2, convergence rate is compared for different choices of the starting function. It can be seen that the convergence is much better if the staring function is below the true solution. However, all three choices have a uniform linear convergence. In Figure 3, the starting function is partly below and partly above the true solution. Convergence for different overlapping sizes is shown. In order to reach a given accuracy, it was observed that the iteration number is reduced by a factor of 2 if we increase the overlapping size by a factor of 2.

6.2 Experiments with the two-level method

Due to the coarse mesh correction, the computed solutions will not increase or decrease monotonically. The first thing we want to show is that the algorithms will not converge if u and ψ are not decomposed properly. We decompose u and ψ as

$$\begin{aligned} u = u_0 + \textstyle\sum_{i=1}^m u_i, \ u_0 = I_H u, \ u_i = I_h(\theta_i(u - u_0)), \\ \psi = \psi_0 + \textstyle\sum_{i=1}^m \psi_i, \ \psi_0 = I_H \psi, \ \psi_i = I_h(\theta_i(\psi - \psi_0)), \end{aligned} \tag{61}$$

i.e. the coarse mesh functions u_0 and ψ_0 are the coarse mesh interpolations for u and ψ respectively. With such an decomposition, the algorithms are not convergent.

The second decomposition we have tried is:

$$\begin{aligned} u = u_0 + \textstyle\sum_{i=1}^m u_i, \ u_0 = 0, \ u_i = I_h(\theta_i u), \\ \psi = \psi_0 + \textstyle\sum_{i=1}^m \psi_i, \ \psi_0 = 0, \ \psi_i = I_h(\theta_i \psi), \end{aligned} \tag{62}$$

i.e. the coarse mesh functions u_0 and ψ_0 are taken to be zero functions. With such a decomposition, it can be proven that the estimate for C_1 is the same as

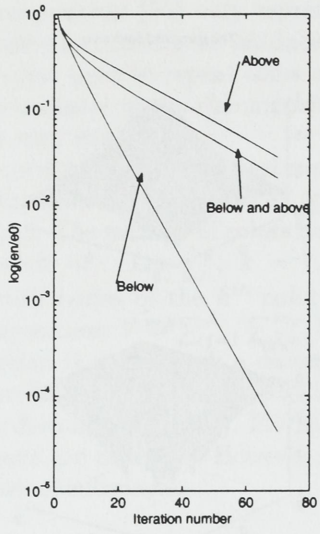

Figure 2: Convergence for the domain decomposition method without the coarse mesh when the starting function is below, above or partly below and partly above the true solution. $h = 4/128$, $H = 4/8$, $\delta = 2h$.

without using the coarse mesh in the decomposition. In the numerical tests, the convergence rate for this decomposition is the same as the domain decomposition method without the coarse mesh, see Figure 4.

Let ψ_0 to be an arbitrary coarse mesh function from V_0. We then decompose ψ as $\psi = \psi_0 + \sum_{i=1}^{m} \psi_i$ with $\psi_i = I_h(\theta_i(\psi - \psi_0))$. The decomposition for u should be taken as in (49) and (51). The analysis indicates that linear convergence shall be obtained for any $\psi_0 \in V_0$. This is in fact observed in the experiments.

In Figure 4, the convergence for different decompositions is compared. The first curve, counting from the top to the bottom, shows the convergence for decomposition (61). It is not convergent. The second curve shows the convergence for the domain decomposition method without the coarse mesh with overlapping size $\delta = 2h$. The third curve shows the convergence with the coarse mesh and with the decomposition given by (62) when the overlapping size is $\delta = 2h$. The convergence is the same as without using the coarse mesh. The last curve shows the convergence with the correct decomposition given by (49) and (51).

In Figure 5, the fine and the coarse meshes are fixed. The convergence for different overlapping sizes is shown. The convergence is better with bigger overlapping sizes.

6.3 Experiments with the multigrid method

For the multigrid method, there are infinitely many choices to decompose $\psi = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \psi_j^i$. For any of these decompositions, the convergence rate is the same just if we decompose u as given in (59). One of the decompositions for ψ is to take $\psi_j^i = 0$ for any $j < J$ and $\psi_J^i = \psi(x_J^i)\phi_J^i(x)$ for $i = 1, 2, \cdots, n_J$, i.e. all the coarser mesh obstacle functions are taken to be zero and only the obstacle on the finest mesh is nonzero. We always start with u being the global obstacle. Convergence for different J is shown in Figure 6. It can be seen that the convergence rate increase slightly with bigger J. For $J = 5$, the rate is 0.78. For $J = 6$, the rate is 0.8. For $J = 7$, the rate is 0.81. For $J = 8$, the rate is 0.85.

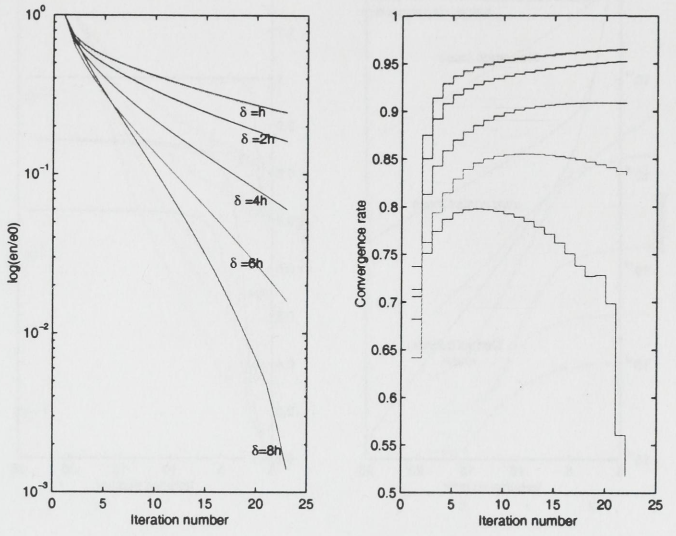

Figure 3: Convergence for the domain decomposition without the coarse mesh when the starting function is partly below and partly above the true solution. $h = 4/128$, $H = 4/8$.

There are some tricks that enable us to compute the decomposition of u given in (58) and (59) very efficiently. For any $v \in S_h$ and $v \geq 0$, we use a vector z_j to store the values $\min_{\tau_j^i} v$ for all the elements $\tau_j^i \subset \mathcal{T}_{h_j}$. As the meshes are nested, the vectors z_j can be computed recursively starting from the finest mesh and ending with the coarsest mesh. From the vectors z_j, it is easy to compute $I_{h_j}^{\ominus} v$ on each level. The value of $I_{h_j}^{\ominus} v$ at a given node is just the smallest value of z_j in the neighboring elements.

7 Conclusion

The decomposition of the obstacle and u can be done very efficiently with the nonlinear operator $I_H^{\ominus}$ for the two-level and the multigrid methods. The complexity of the code is nearly the same as the unconstrained linear case. However, the convergence rate can be improved if we use other interpolation operators. There are many other nonlinear interpolation operators satisfying the properties (34), (35), (36), (37) and (38). Some of these operators satisfy (37) and (38) with a much smaller constant C. The corresponding C_1 for these operators will be much smaller for the two-level and multigrid method. From Theorem 1, the convergence rate with these interpolation operators can be better. This is confirmed in our numerical tests. Extra costs are involved with these operators and there are several alternatives. Intensive numerical tests with these operators will be reported elsewhere.

In condition (3), the nonlinear operator F' is required to be coercive. Condition (8) implies that F' is Lipschitz continuous. The convergence of Theorem 1 can be extended to nonlinear problems under weaker conditions as in [57]. Just assuming

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

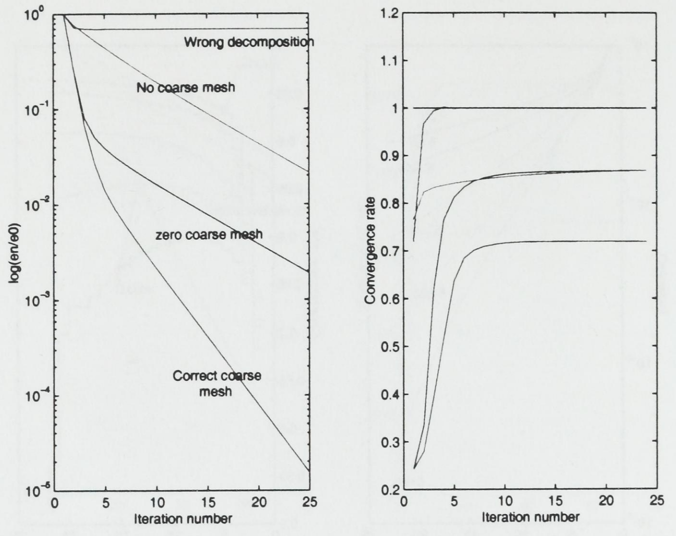

Figure 4: Convergence for different decompositions. $h = 4/128$, $H = 4/8$, $\delta = 2h$.

with some $\kappa > 0, \ell >, p > 1$ and $q > 1$, it can be proved that

$$F(w) - F(u^*) \leq \left(1 - \frac{1}{c_0}\right)(F(u) - F(u^*)) \quad \text{if} \quad p = q.$$

and

$$F(w) - F(u^*) \leq \frac{F(u) - F(u^*)}{\left[1 + c_0 \left|F(u) - F(u^*)\right|^{r-1}\right]^{\frac{1}{r-1}}}, \; r = \frac{p(p-1)}{q(q-1)}, \quad \text{if} \quad p > q.$$

In order to get the above estimates, conditions (7) and (8) also need to be modified correspondingly and can be shown to be valid for all the decompositions given in sections 5.2, 5.3 and 5.4. The constant $c_0 > 1$ is given explicitly as a function of $\alpha, \kappa, \ell, p, q, C_1$ and C_2.

Acknowledgement: The author would like to thank Professor Paul Tseng for careful proof reading and pointing out a mistake in an earlier draft and B. Heimsund for some assistance with the numerical experiments. Thanks are also extended to Professors M. Espedal, Ragnar Winther and J. Xu for insightful discussions and comments.

References

[1] G. E. Alefeld, X. Chen, and F. A. Potra. Numerical validation of solutions of linear complementarity problems. *Numer. Math.*, 83(1):1–23, 1999.

[2] L. Badea. On the schwarz alternating method with more than two subdomains for nonlinear monotone problems. *SIAM J. Numer. Anal.*, 28:179–204, 1991.

[3] L. Badea and J. Wang. Domain decomposition methods for variational inequalities. In J. Wang et al., editor, *Iterative methods in scientific computation*, volume 4 of *Series in computational and applied mathematics*, pages 199–204. IMACS, 1998.

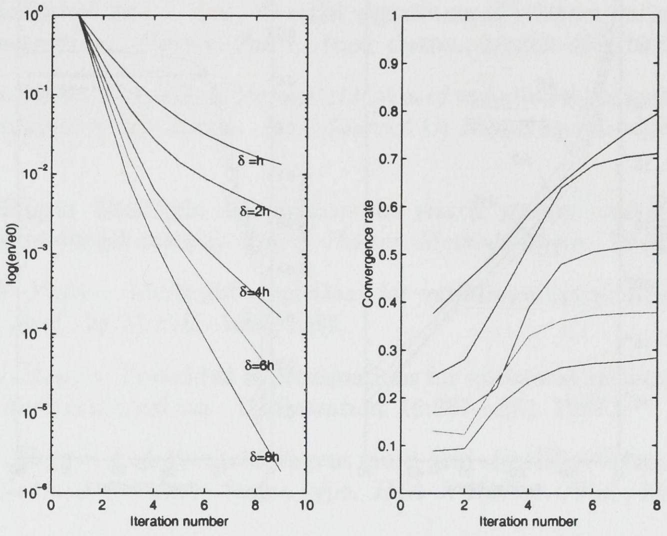

Figure 5: Convergence for the two-level method for decomposition (49) and (51) with different overlaps. $h = 4/128$, $H = 4/8$.

[4] Lori Badea and Junping Wang. An additive Schwarz method for variational inequalities. *Math. Comp.*, 69(232):1341–1354, 2000.

[5] C. Baicci and A. Capelo. *Variational and quasivariational inequalities.* Wiley, Chichester, 1984.

[6] F. Bornemann and H. Yserentant. A basic norm equivalence for the theory of multilevel methods. *Numer. Math.*, 64(4):455–476, 1993.

[7] J. H. Bramble and S. R. Hilbert. Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. *SIAM J. Numer. Anal.*, 7:112–124, 1970.

[8] J. H. Bramble and S. R. Hilbert. Bounds for a class of linear functionals with applications to Hermite interpolation. *Numer. Math.*, 16:362–369, 1970/1971.

[9] 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.

[10] J. H. Bramble, J. E. Pasciak, and J.-C. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55:1–22, 1990.

[11] A. Brandt and C. W. Cryer. Multigrid algorithms for the solution of linear complementary problems arising from free boundary problems. *SIAM J. Sci. Comput.*, 4:655–684, 1983.

[12] J. Cea. *Optimisation – thérie et algorithmes.* Dunod, 1971.

[13] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam, 1978.

[14] R. W. Cottle, J. S. Pang, and R. E. Stone. *The linear complementary problem.* Academic Press, Boston, 1992.

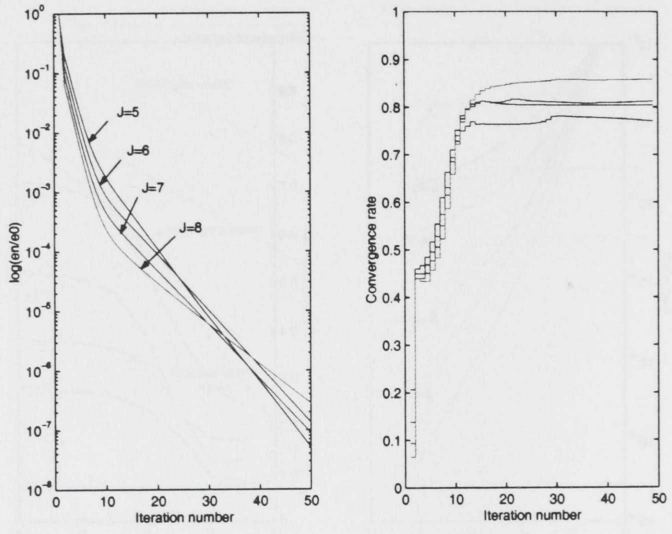

Figure 6: Convergence for the multigrid method

[15] M. Dryja and O. B. Widlund. Domain decomposition algorithms with small overlap. *SIAM J. Sci. Comp.*, 15:604–620, 1994.

[16] G. Duvaut and J. L. Lions. *Inequalities in mechanics and physics.* Springer, Berlin, 1976.

[17] I. Ekeland and R. Temam. *Convex analysis and variational problems.* North–Holland, Amsterdam, 1976.

[18] C. M. Elliot and J. R. Ockendon. *Weak and variational methods for moving boundary problems*, volume 53 of *Research notes in mathematics.* Pitman, London, 1982.

[19] B. Erdmann, M. Frei, R. H. W. Hoppe, R. Kornhuber, and U. Wiest. Adaptive finite element methods for variational inequalities. *East-West J. Numer. Math.*, 1(3):165–197, 1993.

[20] Andreas Frommer and Rosemary A. Renaut. A unified approach to parallel space decomposition methods. *J. Comput. Appl. Math.*, 110(1):205–223, 1999.

[21] Masao Fukushima, Zhi-Quan Luo, and Jong-Shi Pang. A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints. *Comput. Optim. Appl.*, 10(1):5–34, 1998.

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

[23] R. Glowinski. *Numerical methods for nonlinear variational problems.* Springer-Verlag, 1984.

[24] R. Glowinski, J. L. Lions, and R. Trémoléres. *Numerical analysis of variational inequalities.* North-Holland, 1984.

[25] W. Hackbusch and H. D. Mittelmann. On multilevel methods for variational inequalities. *Numer. Math.*, 42:65–76, 1983.

[26] K. H. Hoffmand and J. Zou. Parallel algorithms of schwarz variant for variational inequalities. *Numer. Funct. Anal. Optim.*, 13:449–462, 1992.

[27] K. H. Hoffmann and J. Zou. Parallel solution of variational inequality problems with nonlinear source terms. *IMA Journal Of Numerical Analysis*, 16:31–45, 1996.

[28] R. H. Hoppe. Multigrid solutions to the elastic plastic torsion problem in multiply connected domains. *Int. J. Numer. Methods Engin.*, 26:631–646, 1988.

[29] R. H. W. Hoppe. Multigrid algorithms for variational inequalities. *SIAM J. Numer. Anal.*, 24(5):1046–1065, 1987.

[30] R. H. W. Hoppe. Two-sided approximations for unilateral variational inequalities by multigrid methods. *Optimization*, 16:867—881, 1987.

[31] R. H. W. Hoppe. A globally convergent multi-grid algorithm for moving boundary problems of two-phase Stefan type. *IMA J. Numer. Anal.*, 13(2):235–253, 1993.

[32] R. H. W. Hoppe and R. Kornhuber. Adaptive multilevel methods for obstacle problems. *SIAM J. Numer. Anal.*, 31:301–323, 1994.

[33] D. Kinderlehrer and G. Stampacchia. *An introduction to variational inequalities and their applications.* Academic Press, New York, 1994.

[34] R. Kornhuber. Monotone multigrid methods for elliptic variational inequalities I. *Numer. Math*, 69:167–184, 1994.

[35] R. Kornhuber. Monotone iterations for elliptic variational inequalities, II. *Numer. Math.*, 72:481–499, 1996.

[36] R. Kornhuber. *Adaptive monotone multigrid methods for nonlinear variational problems.* Teubner, Stuttgart, 1997.

[37] Yu. A. Kuznetsov, P. Neittaanmäki, and P. Tarvainen. Schwarz methods for obstacle problems with convection-diffusion operators. In *Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993)*, pages 251–256. Amer. Math. Soc., Providence, RI, 1994.

[38] 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.

[39] T. Lu, T. M. Shih, and C. B. Liem. Parallel algorithm for variational inequalities based on domain decomposition. *Systems Sci. Math. Sci.*, 4:341–348, 1991.

[40] Z.-Q. Luo and P. Tseng. On the linear convergence of descent methods for convex essentially smooth minimization. *SIAM J. Control Optim.*, 30:408–425, 1992.

[41] Z.-Q. Luo and P. Tseng. Error bounds and convergence analysis of feasible descent methods: A general approach. *Ann. Oper. Res.*, 46:157–178, 1993.

[42] J. Mandel. Algebraic study of a multigrid method for some free boundary problems. *Comptes Rendus Academic of Science Paris, Series I*, 298:77–95, 1984.

[43] J. Mandel. A multilevel iterative method for symmetric, positive definite linear complementary problems. *Appl. Math. Optim.*, 11:77–95, 1984.

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

[45] P. Oswald. On discrete norm estimate related to multilevel preconditioners in the finite lelement method. In Ivanov K. G. and Sendov B., editors, *Constructive theory of functions*, Proc. Int. Conf. Varna, 1991, pages 203–214. Bulg. Acad. Sci, 1992.

[46] R. Rodrigues. *Obstacle problems in mathematical physics*. North-Holland, Amsterdam, 1987.

[47] I. Sharapov. Multilevel subspace correction for large scale optimization problems. Technical Report CAM-97-31, UCLA, computational and applied mathematics, 1997.

[48] M. V. Solodov and P. Tseng. Modified projection type methods for monotone variational inequalities. *SIAM J. Control Optim.*, 34:1814–1830, 1996.

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

[50] 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, 1994.

[51] 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 at http://www.mi.uib.no/~tai.

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

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

[54] X.-C. Tai. Convergence rate analysis of domain decomposition methods for obstacle problems. Technical Report 124, Department of Mathematics, University of Bergen, January, 1999.

[55] 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.

[56] X.-C. Tai and P. Tseng. Convergence rate analysis of an asynchronous space decomposition method for convex minimization. *Math. Comput.*, to appear. Also as Report-120, Department of Mathematics, University of Bergen, August, 1998, available at http://www.mi.uib.no/~tai.

[57] X-C. Tai and J.-C. Xu. Global convergence of subspace correction methods for convex optimization problems. *Math. Comput.*, to appear. Also as Report-114, Department of Mathematics, University of Bergen, August, 1998, available at http://www.mi.uib.no/~tai.

[58] Xue-Cheng Tai. Parallel function decomposition methods and numerical applications. In J. Wang, M. B. Allen, B. M. Chen, and T. Mathew, editors, *Iterative methods in scientific computation*, volume 4 of *Series in computational and applied mathematics*, pages 231–238. IMACS, 1998.

[59] P. Tarvainen. Two-level Schwarz method for unilateral variational inequalities. *IMA J. Numer. Anal.*, 14:1–18, 1998.

[60] O. 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.

[61] J. Xu. Iteration methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.

[62] J. Zeng and S. Zhou. On monotone and geometric convergence of schwarz methods for two sided obstacle problems. *SIAM J. Numer. Anal.*, 35:600–616, 1998.