38

An Asynchronous Space Decomposition Method

Xue-Cheng Tai, [1] Paul Tseng [2]

Introduction

With the advent of multiprocessor computing systems, there has been much work in the design and analysis of iterative methods that can take advantage of the parallelism to solve large algebraic problems. In these methods, the computation per iteration is distributed over the processors and each processor communicates the result of its computation to the other processors, possibly subject to communication or computation delays. By using a model of asynchronous computation proposed by Chazan and Miranker [CM69], these methods have been analyzed quite extensively (see [BT89]and references therein). However, aside from the easy case where the algorithmic mapping is a contraction with respect to the L^∞-norm, there has been few studies of the convergence rate of these methods.

In this paper, we study the convergence rate of asynchronous block Jacobi and block Gauss-Seidel methods for finite or infinite dimensional convex minimization of

1 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.
2 Department of Mathematics, Box 354350, University of Washington, Seattle, WA 98195, U.S.A. Email: tseng@math.washington.edu and URL: http://www.math.washington.edu/~tseng.
The work of the first author was supported by the Engineering and Physical Sciences Research Council, United Kingdom, Grant Reference GR/L98862. The work of the second author was supported by the National Science Foundation, Grant No. CCR-9311621.
Eleventh International Conference on Domain Decomposition Methods
Editors Choi-Hong Lai, Petter E. Bjørstad, Mark Cross and Olof B. Widlund

the form

$$\min_{v_i \in K_i, i=1,\ldots,m} F\left(\sum_{i=1}^{m} v_i\right), \tag{1}$$

where each K_i is a nonempty closed convex set in a real reflexive Banach space V and F is a real-valued lower semicontinuous Gâteau-differentiable function that is strongly convex on $\sum_{i=1}^{m} K_i$. Our interest in these methods stems from their close connection to relaxation methods for nonlinear network flow (see [BCEZ95], [BT89], [TBT90] and references therein) and to domain decomposition (DD) and multigrid (MG) methods for solving elliptic partial differential equations (see [CM94], [DH97], [SBG96], [TE98], [TX98], [Xu92] and references therein). For example, the additive and the multiplicative Schwarz methods may be viewed as block Jacobi and block Gauss-Seidel methods applied to linear elliptic partial differential equations reformulated as (1). DD and MG methods are also useful as preconditioners and it can be shown that such preconditioning improves the condition number of the discrete approximation [BZar], [CM94], [SBG96], [Xu92]. In addition, DD and MG methods are well suited for parallel implementation, for which both synchronous and asynchronous versions have been proposed. Of the work on asynchronous methods [BMPSar], [BMPS95], [FSS97], [McC89], we especially mention the numerical tests by Frommer et al. [FSS97] which showed that, through improved load balancing, asynchronous methods can be advantageous in solving even simple linear equations. Although these tests did not use the coarse mesh in its implementation of the DD method, it is plausible that the asynchronous method would still be advantageous when the coarse mesh is used. An important issue concerns the convergence and convergence rate of the above methods. In the case where the equation is linear (corresponding to F being quadratic and $K_1, ..., K_m$ being suitable subspaces of V) or almost linear, this has been much studied for synchronous methods such as block Jacobi and block Gauss-Seidel methods (see [BZar], [CM94], [SBG96], [Xu92] and references therein) but little studied for asynchronous methods [BMPSar], [BMPS95], [McC89]. In the case where the equation is nonlinear (corresponding to $K_1, ..., K_m$ being suitable subspaces of V), there are some convergence studies for synchronous methods [CSar], [DH97], [Sha97], [TE98], [TX98], and none for asynchronous methods. In the case where $K_1, ..., K_m$ are not all subspaces, there are some convergence studies for synchronous methods and, in particular, block Jacobi and Gauss-Seidel methods (see [Kor97], [LT92], [LT93], [Tai92], [Tai95], etc.) but none for asynchronous methods.

Our contributions are two-fold. First, we consider an asynchronous version of block Jacobi and block Gauss-Seidel methods for solving (1), and we show that, under a Lipschitzian assumption on the Gâteau derivative F' and a norm equivalence assumption on the product of $K_1, ..., K_m$ and their sum (see (5) and (6)), this asynchronous method attains global linear rate of convergence with a convergence factor that can be explicitly estimated (see Theorem 1). This provides a unified convergence and convergence rate analysis for such asynchronous methods. Second, we apply the above convergence result to linearly constrained convex programs and, in particular, nonlinear network flow problems. This yields convergence rate results for some asynchronous network relaxation methods (see §38). We also apply the above convergence result to certain nonlinear elliptic partial differential equations. This yields convergence rate results for some parallel DD and MG methods applied to these equations and, in particular, the convergence factor is shown not to depend on the mesh parameters (see §38). Although alternative approaches such as Newton-type methods have been applied to develop synchronous DD and MG methods for nonlinear partial differential equations [BR82], [Bra77], [HR89], [Xu96], these methods use the traditional DD and MG approach or use a special two-grid treatment.

Problem Description and Space Decomposition

Let V be a real reflexive Banach space with norm $\|\cdot\|$ and let V' be its dual space, i.e., the space of all real-valued linear continuous functionals on V. The value of $f \in V'$ at $v \in V$ will be denoted by $\langle f, v\rangle$, i.e., $\langle\cdot,\cdot\rangle$ is the duality pairing of V and V'. We wish to solve the following minimization problem

$$\min_{v\in K} F(v) \ , \tag{2}$$

where K is a nonempty closed (in the strong topology) convex set in V and $F : V \mapsto \Re$ is a lower semicontinuous convex Gâteau-differentiable function. We assume F is strongly convex on K or, equivalently, its Gâteau derivative $\lim_{t\to 0}(F(v + tw) - F(v))/t$, which is a well-defined linear continuous functional of w denoted by $F'(v)$ (so $F' : V \mapsto V'$), is strongly monotone on K, i.e.,

$$\langle F'(u) - F'(v), u - v\rangle \geq \sigma\|u - v\|^2, \quad \forall u, v \in K, \tag{3}$$

where $\sigma > 0$. It is known that, under the above assumptions, (2) has a unique solution $\bar{u}$ [GT89].

We assume that the constraint set K can be decomposed as the Minkowski sum:

$$K = \sum_{i=1}^{m} K_i, \tag{4}$$

for some nonempty closed convex sets K_i in V, $i = 1, ..., m$. This means that, for any $v \in K$, we can find $v_i \in K_i$, not necessarily unique, satisfying $\sum_{i=1}^m v_i = v$ and, conversely, for any $v_i \in K_i$, $i = 1, ..., m$, we have $\sum_{i=1}^m v_i \in K$. Following Xu [Xu92], we call (4) a space decomposition of K, with the term "space" used loosely here. Then we may reformulate (2) as the minimization problem (1), with $(\bar{u}_1, ..., \bar{u}_m)$ being a solution (not necessarily unique) of (1) if and only if $\bar{u}_i \in K_i$ for $i = 1, ..., m$ and $\sum_{i=1}^m \bar{u}_i = \bar{u}$. As was noted earlier, the reformulated problem (1) is of interest because methods such as DD and MG methods may be viewed as block Jacobi and block Gauss-Seidel methods for its solution. The method we study will be an asynchronous version of these methods. The above reformulation was proposed in [Xu92] (for the case where F is quadratic and $K = V$) to give a unified analysis of DD and MG methods for linear elliptic partial differential equations. The general case was treated in [Tai92], [Tai95] (also see [Tai94], [TE98] for the case of $K = V$).

For the above space decomposition, we will assume that there is a constant $C_1 > 0$ such that for any $v_i \in K_i, i = 1, ..., m$, there exists $\bar{u}_i \in K_i$ satisfying

$$\bar{u} = \sum_{i=1}^{m} \bar{u}_i \quad \text{and} \quad \left(\sum_{i=1}^{m} \|\bar{u}_i - v_i\|^2\right)^{\frac{1}{2}} \leq C_1\left\|\bar{u} - \sum_{i=1}^{m} v_i\right\|. \tag{5}$$

See [CM94], [Tai95], [TE98], [Xu92] for similar assumptions. We will also assume F' has a weak Lipschitzian property in the sense that there is a constant $C_2 > 0$ such that

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\langle F'(w_{ij}+u_{ij}) - F'(w_{ij}), v_i\rangle \le C_2\Big(\sum_{j=1}^{m}\max_{i=1,\ldots,m}\|u_{ij}\|^2\Big)^{\frac{1}{2}}\Big(\sum_{i=1}^{m}\|v_i\|^2\Big)^{\frac{1}{2}}, \quad \forall w_{ij}\in K, u_{ij}\in K_j^{\ominus}, v_i\in K_i^{\ominus},\ i,j=1,\ldots,m, \tag{6}$$

where we define the set difference $K_i^{\ominus} = \{u-v : u, v \in K_i\} \subset V$. The above assumption generalizes those in [Tai95], [TE98], [TX98] for the case of K_i being a subspace, for which $K_i^{\ominus} = K_i$.

Furthermore, we will paint each of the sets $K_1, \ldots, K_m$ one of c colors, with the colors numbered from 1 up to c, such that sets painted the same color $k \in \{1, \ldots, c\}$ are orthogonal in the sense that, for all $u \in K$ and $v_i \in K_i^{\ominus}$, $i \in I(k)$,

$$\Big\|\sum_{i\in I(k)} v_i\Big\|^2 = \sum_{i\in I(k)}\|v_i\|^2, \tag{7}$$

$$\Big\langle F'\Big(u + \sum_{i\in I(k)} v_i\Big), \sum_{i\in I(k)} v_i\Big\rangle \le \sum_{i\in I(k)}\langle F'(u+v_i), v_i\rangle, \tag{8}$$

where $I(k) = \{i \in \{1, \ldots, m\} : K_i$ is painted color $k\}$. See [CM94], [TX98] for similar orthogonal decompositions in the case K_i is a subspace. Thus $I(1), \ldots, I(c)$ are disjoint subsets of $\{1, \ldots, m\}$ whose union is $\{1, \ldots, m\}$ and $I(k)$ comprises the indexes of the sets painted the color k. Although $c = m$ is always a valid choice, in some of the applications that we will consider, it is essential that c be independent of m. In the context of a network flow problem, each set K_i may correspond to a node of the network and sets are painted different colors if their corresponding nodes are joined by an arc. In the context of a partial differential equation defined on a domain $\Omega \subset \Re^d$, each set K_i may correspond to a subdomain of Ω and sets are painted different colors if their corresponding subdomains intersect (see §38, §38).

Remark 1: It can be seen that condition (6) is implied by the following strengthened Cauchy-Schwarz inequality (also see [SBG96], [Xu92] for the case of quadratic F and subspace K_i):

$$\langle F'(w_{ij}+u_{ij}) - F'(w_{ij}), v_i\rangle \le \epsilon_{ij}\|u_{ij}\|\|v_i\|, \quad \forall w_{ij}\in K, u_{ij}\in K_j^{\ominus}, v_i \in K_i^{\ominus},$$

with C_2 being the spectral radius of the symmetric matrix $\mathcal{E} = [\epsilon_{ij}]_{i,j=1}^{m}$.

Remark 2: For locally strongly convex problems, the constants σ, C_1, C_2 may depend on u, v, v_i, w_{ij}, u_{ij}. In this case, the subsequent convergence estimate should be viewed as being local in nature, i.e., it is valid when the iterated solutions lie in a neighborhood of the true solution (see §38).

An Asynchronous Space Decomposition Method

Since F is lower semicontinuous and strongly convex, for each $(u_1, ..., u_m) \in K_1 \times \cdots \times K_m$ and each $i \in \{1, ..., m\}$, there exists a unique $w_i \in K_i$ satisfying

$$F\left(\sum_{j\neq i} u_j + w_i\right) \leq F\left(\sum_{j\neq i} u_j + v_i\right), \quad \forall v_i \in K_i \tag{9}$$

(see [GT89]). Let $\pi_i(u_1, ..., u_m)$ denote this w_i. Then $(\pi_1, ..., \pi_m)$ may be viewed as the algorithmic mapping associated with the block Jacobi method for solving (1). Consider an asynchronous version of the block Jacobi method, parameterized by a stepsize $\gamma \in (0, 1]$ which for simplicity we assume to be fixed, that generates a sequence of iterates $(u_1(t), ..., u_m(t))$, $t = 0, 1, ...$, with $(u_1(0), ..., u_m(0)) \in K_1 \times \cdots \times K_m$ given, according to the updating formula:

$$u_i(t+1) \;=\; u_i(t) + \gamma s_i(t), \quad i = 1, ..., m, \tag{10}$$

where we define

$$s_i(t) = w_i(t) - u_i(t) \quad \text{if} \quad t \in T^i \quad \text{and} \quad s_i(t) = 0 \quad \text{otherwise}, \tag{11}$$

$$\text{with} \quad w_i(t) = \pi_i\left(u_1(\tau_1^i(t)), ..., u_m(\tau_m^i(t))\right), \tag{12}$$

and T^i is some subset of $\{0, 1, ...\}$ and each $\tau_j^i(t)$ is some nonnegative integer not exceeding t. Since each K_i is convex and $\gamma \in (0, 1]$, an induction argument shows that $(u_1(t), ..., u_m(t)) \in K_1 \times \cdots \times K_m$ for all $t = 0, 1, ...$

We will assume that the iterates are updated in a *partially asynchronous* manner [BT89], i.e., there exists an integer $B \geq 1$ such that

$$\{t, t+1, ..., t+B-1\} \cap T^i \neq \emptyset \qquad t = 0, 1, ..., \ \forall i, \tag{13}$$

$$0 \leq t - \tau_j^i(t) \leq B - 1 \quad \text{and} \quad \tau_i^i(t) = t \qquad \forall t \in T^i, \ \forall i, j. \tag{14}$$

Remark 3: The above asynchronous method models a situation in which computation is distributed over m processors with the ith processor being responsible for updating u_i and communicating the updated value to the other processors. T^i is the set of "times" at which u_i is updated by processor i (by applying π_i to its current copy of $(u_1, ..., u_m)$); $u_i(t)$ is the value of u_i known to processor i at time t; and $\tau_j^i(t)$ is the time at which the value of u_j used by processor i at time t is generated by processor j, so $t - \tau_j^i(t)$ is the communication delay from processor j to processor i at time t. Thus, the processors need not wait for each other when updating $(u_i)_{i=1}^m$, and the values used in the computation may be out-of-date.

Convergence Rate of the Asynchronous Method

Below is our main convergence result, showing that the iterates $(u_1(t), ..., u_m(t))$ generated by the asynchronous method (10)–(14) attain linear rate of convergence, with a factor that depends on σ, C_1, C_2, c and B, γ only. We refer the proof to [TT98].

Theorem 1 *Consider the minimization problem* (2) *and the space decomposition* (4) *of §38 (see* (3), (5)–(8)*). Let* $(u_1(t), ..., u_m(t))$, $t = 0, 1, ...,$ *be generated by the asynchronous space decomposition method of §38 (see* (10)–(12) *and* (13), (14)*). Denote* $u(t) = \sum_{j=1}^{m} u_j(t)$*. Then, there exist* $\gamma_0 \in (0, 1)$ *and* $\varrho \in (0, 1)$*, depending on* σ, C_1, C_2, c *and* B *only, such that when* $\gamma \leq \gamma_0$*, there holds*

$$F(u(nB)) - F(\bar{u}) \leq \varrho^{n-1} \max\Big\{F(u(B)) - F(\bar{u}), \sum_{t=0}^{B-1} \sum_{i=1}^{m} \|s_i(t)\|^2\Big\}, \quad n = 1, 2, ...,$$

where $\bar{u}$ *denotes the unique solution of* (2)*. Moreover,* $u(t)$ *converges strongly to* $\bar{u}$ *and, for each* $i \in \{1, ..., m\}$*,* $u_i(t)$ *converges strongly as* $t \to \infty$*.*

Applications to Convex Programming

Primal Applications

Consider the case of the problem (2), where $V = V' = \Re^n$, $F : \Re^n \mapsto \Re$ is a differentiable convex function, and K is a nonempty polyhedral set in $\Re^n$. Then F is continuous [Roc70] and continuously differentiable [Roc70]. We assume that the gradient $F' = (\frac{\partial F}{\partial x_j})_{j=1}^{n}$ is strongly monotone and Lipschitz continuous on K and we choose a space decomposition (4) such that each K_i is a polyhedral set.

Since each K_i is a polyhedral set, a Lipschitzian property of the solution set of a linear system (see [BT96]) implies that, for any $v_i \in K_i$, $i = 1, ..., m$, there exists $\bar{u}_i \in K_i$ satisfying (5), where C_1 depends on m and certain condition numbers for K_i, $i = 1, ..., m$. In cases where each K_i has a simple structure, C_1 may be estimated explicitly. Also, an analysis similar to that used for (21) shows that (6) holds with $C_2 = L\hat{c}$, where L is the Lipschitz constant for F' and $\hat{c}$ is the maximum number of sets K_j that are not orthogonal to an arbitrary set K_i. To color the sets such that (7)–(8) hold, it suffices to paint K_i and K_j different colors whenever they are not orthogonal, i.e., $(v_i)^T v_j \neq 0$ for some $v_i \in K_i, v_j \in K_j$.

Dual Applications

Consider the linearly constrained convex program

$$\text{minimize} \quad G(x) \quad \text{subject to} \quad Ax = b, \tag{15}$$

where $G : \Re^n \mapsto \Re$ is a strictly convex differentiable function, $b \in \Re^m$, and $A \in \Re^{m\times n}$ has nonzero rows. We assume there exists $\tilde{x} \in \Re^n$ satisfying $A\tilde{x} = b$. By attaching Lagrange multipliers $\lambda \in \Re^m$ to the equations $Ax = b$ in (15), we obtain the Lagrangian dual problem:

$$\min_{\lambda\in\Re^m} \; G^*(A^T\lambda) - b^T\lambda, \tag{16}$$

where G^* is the convex conjugate of G defined by (see [GT89], [Roc70])

$$G^*(u) = \sup_{x\in\Re^n} \left\{u^T x - G(x)\right\}.$$

The convex programs (15) and (16) are dual in the sense that one has a solution if and only if the other does and these solutions satisfy $G'(x) = A^T\lambda$ [Roc70]. Using $b = A\tilde{x}$, we can rewrite the dual problem (16) in the form of (2) with

$$F(u) = G^*(u) - \tilde{x}^T u, \qquad K = \{u \in \Re^n : u = A^T\lambda \text{ for some } \lambda \in \Re^m\}. \tag{17}$$

We assume that $(G^*)'$ is strongly monotone and Lipschitz continuous on $\Re^n$, so that F satisfies (3) for some $\sigma > 0$.

Let $\bar{u}$ denote the unique solution of (2) and let A_i denote the ith row of A. We decompose K in the form (4) with subspaces

$$K_i = \{u_i \in \Re^n \ : \ u_i = A_i^T \lambda_i \text{ for some } \lambda_i \in \Re\}.$$

It was shown in [TT98] that, for any $v_i \in K_i$, $i = 1, ..., m$, there exists $\bar{u}_i \in K_i$ satisfying (5), where C_1 depends on A only. It was also shown that (6) holds with $C_2 = L\hat{c}$, where L is the Lipschitz constant for $(G^*)'$ and $\hat{c}$ is the maximum number of rows A_j that are not orthogonal to an arbitrary row A_i. Since two subspaces K_i and K_j are orthogonal if and only if $A_i A_j^T = 0$, we can color $K_1, ..., K_m$ as discussed in §38 so that (7)–(8) hold.

Applications to Partial Differential Equations

The first partial differential equation corresponds to the minimization problem (2) with $V = K = H_0^1(\Omega)$ and

$$\langle F'(u), v\rangle = \int_\Omega \left(\sum_{i=1}^d a_i(x, u, \nabla u)\partial_i v + a_0(x, u, \nabla u)v - fv \right) dx, \tag{18}$$

where Ω is a suitable domain of $\Re^d$, $f \in L^2(\Omega)$, and the nonlinear coefficient a_i, $i = 0, 1, \cdots d$ are such that (3) is satisfied for some $\sigma > 0$ (see [TT98]).

The second partial differential equation corresponds to the minimization problem (2) with $V = K = H_0^1(\Omega)$ and

$$F(v) = \int_\Omega \left(\frac{1}{2}|\nabla v|^2 + \frac{1}{4}v^4 - fv \right) dx, \tag{19}$$

with Ω, f as above and with $d \in \{2, 3\}$. The corresponding equation is the simplified Ginzburg-Landau equation for superconductivity:

$$-\Delta u + u^3 = f \quad \text{in} \quad \Omega, \qquad \text{and} \qquad u = 0 \quad \text{on} \quad \partial\Omega, \tag{20}$$

where u is the wave function, which is valid in the absence of internal magnetic field. It can be shown that this F satisfies (3) for some $\sigma > 0$ (see [TT98]).

Domain decomposition methods

In DD methods, the domain Ω is decomposed into the disjoint union of subdomains Ω_i, $i = 1, ..., m$, and their boundary, i.e., $\Omega \cup \partial\Omega = \cup_{i=1}^m (\Omega_i \cup \partial\Omega_i)$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$.

The subdomains, which are assumed to form a regular quasi-uniform division (see p. 124 and Eq. (3.2.28) of [Cia78] for definitions) with a specified maximum diameter of H, are the finite elements of the coarse mesh. To form the fine mesh for the finite element approximations, we further divide each Ω_i into finite elements of size h such that all the fine-mesh elements together form a regular finite element division of Ω. We denote this fine division by $\mathcal{T}_h$. For each Ω_i, we consider an enlarged subdomain $\Omega_i^\delta = \{e \in \mathcal{T}_h : \text{dist}(e, \Omega_i) \leq \delta\}$, where $\text{dist}(e, \Omega_i) = \min_{x \in e, y \in \Omega_i} |x - y|$. The union of Ω_i^δ, $i = 1, ..., m$, covers Ω with overlap proportional to δ. Let $K_0 \subset H_0^1(\Omega)$ and $K \subset H_0^1(\Omega)$ denote the continuous, piecewise rth-order polynomial ($r \geq 1$) finite element subspaces, with zero trace on $\partial\Omega$, over the H-level and h-level subdivisions of Ω respectively. For $i = 1, ..., m$, let K_i denote the continuous, piecewise rth-order polynomial finite element subspace with zero trace on the boundary $\partial\Omega_i^\delta$ and extended to have zero value outside $\Omega_i^\delta \cup \partial\Omega_i^\delta$. Then $K_i^\ominus = K_i$ for $i = 0, 1, ..., m$, and it can be shown that the space decomposition (4), with summation index from 0 to m, holds. We assume that the overlapping subdomains are chosen such that each subdomain Ω_i^δ and its corresponding finite element subspace K_i can be painted one of n_c colors (numbered from 1 to n_c), with subdomains painted the same color being pairwise non-intersecting. The coarse mesh and its corresponding subspace K_0 are painted the color 0. Moreover, n_c should be independent of h. For general domain Ω, finding overlapping subdomains with such property is nontrivial. If Ω is the Cartesian product of intervals, we can easily find overlapping subdomains with $n_c = 2$ if $d = 1$, and $n_c \leq 4$ if $d = 2$, and $n_c \leq 6$ if $d = 3$.Then the total number of colors needed for (7) and (8) to hold is $c = n_c + 1$.

Let $\{\theta_i\}_{i=1}^m$ be a smooth partition of unity with respect to $\{\Omega_i\}_{i=1}^m$, i.e., $\theta_i \in C_0^\infty(\Omega)$ with $\theta_i \geq 0$, $\theta_i = 0$ outside of Ω_i, and $\sum_{i=1}^m \theta_i = 1$. Let I_h be the finite element interpolation mapping onto K which uses the function values at the h-level nodes. For any $v \in K$, let v_0 be the projection in the L^2-norm of v onto K_0, i.e., $v_0 \in K_0$ and $\int_\Omega (v_0 - v)\phi \, dx = 0$ for all $\phi \in K_0$, and let $v_i = I_h(\theta_i(v - v_0))$. Then, it can be seen that $v_i \in K_i$ for $i = 0, 1, ..., m$ and satisfy $v = \sum_{i=0}^m v_i$ [SBG96], [Xu92]. By further choosing θ_i so that $|\nabla\theta_i|$ has a certain boundedness property, it was shown in [TX98] that for any $v_i \in K_i$, $i = 0, 1, ..., m$, there exists $\bar{u}_i \in K_i$ satisfying (5) (with summation index from 0 to m), where C is independent of m and the mesh parameters, and

$$C_1 = C\sqrt{c}\left(1 + \left(\frac{H}{\delta}\right)^{\frac{1}{2}}\right).$$

For F given by (18) or (19), it was shown in [TT98] that

$$\begin{aligned} \sum_{i=0}^m \sum_{j=0}^m \langle F'(w_{ij} + u_{ij}) &- F'(w_{ij}), v_i \rangle \leq \tilde{C}_2 \Big(\sum_{j=0}^m \max_{i=0,1,\ldots,m} \|u_{ij}\|^2 \Big)^{\frac{1}{2}} \Big(\sum_{i=0}^m \|v_i\|^2 \Big)^{\frac{1}{2}} \\ &+ (1 + C) \Big(\sum_{i=1}^m \|u_{i0}\|_{H^1(\Omega_i^\delta)}^2 \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^m \|v_i\|^2 \Big)^{\frac{1}{2}}. \end{aligned} \quad (21)$$

with $\tilde{C}_2$ a constant depending on $C\alpha, c, \hat{c}$ only. Compared with (6) (with $i, j = 0, 1, ..., m$), we see that (21) has an extra term on the right-hand side. It was shown in [TT98] that this extra term does not affect the convergence rate result of §38.

Multigrid methods

In MG methods, Ω is divided into a finite element triangulation $\mathcal{T}$ by a successive refinement process. More precisely, we have $\mathcal{T} = \mathcal{T}_J$ for some $J > 1$, where $\mathcal{T}_k$, $k = 1, ..., J$, is a nested sequence of regular quasi-uniform triangulation, i.e., $\mathcal{T}_k$ is a collection of simplexes $\mathcal{T}_k = \{\tau_i^k\}$ of size (i.e., maximum diameter) h_k such that $\Omega = \cup_i \tau_i^k$ and for which the quasi-uniformity constants are independent of k [Cia78] and with each simplex in $\mathcal{T}_{k-1}$ being the union of simplexes in $\mathcal{T}_k$. We further assume that there is a constant $r < 1$, independent of k, such that h_k is proportional to r^{2k}.

Corresponding to each triangulation $\mathcal{T}_k$, we define the finite element subspace:

$$\mathcal{M}_k = \{v \in H_0^1(\Omega) \ : \quad v|_\tau \in \mathcal{P}_1(\tau), \quad \forall\, \tau \in \mathcal{T}_k\},$$

where $\mathcal{P}_1(\tau)$ denotes the space of real-valued linear functions of d real variables defined on τ. We associate with $\mathcal{M}_k$ a nodal basis, denoted by $\{\phi_i^k\}_{i=1}^{n_k}$, that satisfies $\phi_i^k \in \mathcal{M}_k$ and

$$\phi_i^k(x_j^k) = \delta_{ij}, \quad \text{the Kronecker function},$$

where $\{x_i^k\}_{i=1}^{n_k}$ is the set of all interior nodes of the triangulation $\mathcal{T}_k$. For each such nodal basis function, we define the one-dimensional subspace: $K_i^k = \text{span}\ (\phi_i^k)$. Then, $(K_i^k)^\ominus = K_i^k$ and we have the following space decomposition:

$$K = \sum_{k=1}^{J} \sum_{i=1}^{n_k} K_i^k \quad \text{with} \quad K = \mathcal{M}_J.$$

On each level k, we color the nodes of $\mathcal{T}_k$ so that neighboring nodes are always of a different color. The number of colors needed for a regular mesh is a constant independent of the mesh parameters, which we denote by n_c. Then the total number of colors needed for (7) and (8) (with summation indices adjusted accordingly) to hold is $c = n_c J$. Also, it can be shown that (6) holds and that, for any $v_i^k \in K_i^k$, $i = 1, ..., n_k$, $k = 1, ..., J$, there exists $\bar{u}_i^k \in K_i^k$ satisfying (5) (with summation indices adjusted accordingly), where C_1 and C_2 do not depend on h and J.

REFERENCES

[BCEZ95] Bertsekas D. P., Castañon D., Eckstein J., and Zenios S. A. (1995) Parallel computing in network optimization. In Ball M. O., Magnanti T. L., Monma C. L., and Nemhauser G. L. (eds) *Handbooks in Operations Research and Management Science 7: Network Models*, pages 331–399. Elsevier Scientific, Amsterdam.

[BMPS95] Bru R., Migallón V., Penadés J., and Szyld D. B. (1995) Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Trans. Numer. Anal.* 3: 24–38.

[BMPSar] Bai Z., Migallón V., Penadés J., and Szyld D. B. (to appear) Block and asychronous two-stage methods for mildly nonlinear systems. *Numer. Math.* .

[BR82] Bank R. and Rose D. J. (1982) Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.* 39: 453–465.

[Bra77] Brandt A. (1977) Multilevel adaptive solutions to boundary value problems. *Math. Comp.* 31: 333–309.

[BT89] Bertsekas D. P. and Tsitsiklis J. N. (1989) *Parallel and Distributed Computation: Numerical Methods.* Prentice-Hall, Englewood Cliffs.

[BT96] Burke J. V. and Tseng P. (1996) A unified analysis of Hoffman's bound via Fenchel duality. *SIAM J. Optim.* 6: 265–282.

[BZar] Bramble J. and Zhang X. (to appear) *The Analysis of Multigrid Methods.* North-Holland.

[Cia78] Ciarlet P. G. (1978) *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam.

[CM69] Chazan D. and Miranker W. L. (1969) Chaotic relaxation. *Lin. Algeb. Appl.* 2: 199–222.

[CM94] Chan T. F. and Mathew T. P. (1994) Domain decomposition algorithms. *Acta Numerica* pages 61–143.

[CSar] Chan T. F. and Sharapov I. (to appear) Subspace correction multilevel methods for elliptic eigenvalue problems. In Bjørstad P., Espedal M., and Keyes D. (eds) *Proceedings of the 9th International Domain Decomposition Conference.* John Wiley and Sons.

[DH97] Dryja M. and Hackbusch W. (1997) On the nonlinear domain decomposition method. *BIT* 37: 296–311.

[FSS97] Frommer A., Schwardt H., and Szyld D. B. (1997) Asychronous weighted additive Schwarz methods. *Electronic Trans. Numer. Anal.* 5: 48–61.

[GT89] Glowinski R. and Tallec P. L. (1989) *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics.* SIAM, Philadelphia.

[HR89] Hackbusch W. and Reusken A. (1989) Analysis of a damped nonlinear multilevel method. *Numer. Math.* 55: 225–246.

[Kor97] Kornhuber R. (1997) *Adaptive monotone multigrid methods for nonlinear variational problems.* Teubner, Stuttgart.

[LT92] Luo Z.-Q. and Tseng P. (1992) On the linear convergence of descent methods for convex essentially smooth minimization. *SIAM J. Control Optim.* 30: 408–425.

[LT93] Luo Z.-Q. and Tseng P. (1993) Error bounds and convergence analysis of feasible descent methods: A general approach. *Ann. Oper. Res.* 46: 157–178.

[McC89] McCormick S. F. (1989) *Multilevel Adaptive Methods for Partial Differential Equations*, volume 6 of *Frontiers in Applied Mathematics.* SIAM, Philadelphia.

[Roc70] Rockafellar R. T. (1970) *Convex Analysis.* Princeton University Press, Princeton.

[SBG96] Smith B. F., Bjørstad P. E., and Gropp W. D. (1996) *Domain Decomposition: Parallel Multilevel Algorithms for Elliptic Partial Differential Equations.* Cambridge University Press, Cambridge.

[Sha97] Sharapov I. A. (1997) Multilevel subspace correction for large scale optimization problems. Technical Report CAM-97-31, UCLA, Computational and Applied Mathematics.

[Tai92] Tai X.-C. (1992) Parallel function decomposition and space decomposition methods with applications to optimisation, splitting and domain decomposition. *Preprint No. 231–1992, Institut für Mathematik, TechnischeUniversität Graz .*

[Tai94] Tai X.-C. (1994) Parallel function and space decomposition methods. In Neittaanmäki P. (ed) *Finite element methods, fifty years of the Courant element, Lecture notes in pure and applied mathematics*, volume 164, pages 421–432. Marcel Dekker, New York.

[Tai95] Tai X.-C. (1995) Parallel function decomposition and space decomposition methods: Part II. Space decomposition. *Beijing Mathematics* 1, part 2: 135–152.

[TBT90] Tseng P., Bertsekas D. P., and Tsitsiklis J. N. (1990) Partially asynchronous, parallel algorithms for network flow and other problems. *SIAM J. Control Optim.* 28: 678–710.

[TE98] Tai X.-C. and Espedal M. (1998) Rate of convergence of some space decomposition method for linear and nonlinear elliptic problems. *SIAM J. Numer. Anal.* 35: 1558–1570.

[TT98] Tai X.-C. and Tseng P. (August, 1998) Convergence rate analysis of an asynchronous space decomposition method for convex minimization. Technical Report 120, Department of Mathematics, University of Bergen.

[TX98] Tai X.-C. and Xu J.-C. (March, 1998) Global convergence of subspace correction methods for convex optimization problems. Technical Report No.-114, Department of Mathematics, University of Bergen, Norway.

[Xu92] Xu J.-C. (1992) Iteration methods by space decomposition and subspace correction. *SIAM Rev.* 34: 581–613.

[Xu96] Xu J.-C. (1996) Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.* 33: 1759–1777.