

Department
of
APPLIED MATHEMATICS

Convergence Rate Analysis of an Asynchronous Space
Decomposition Method for Convex Minimization

by

Xue-Cheng Tai and Paul Tseng

Report no. 120

August 1998



UNIVERSITY OF BERGEN
Bergen, Norway



Department of Mathematics
University of Bergen
5008 Bergen
Norway

ISSN 0084-778x

Convergence Rate Analysis of an Asynchronous Space Decomposition Method for Convex Minimization

by

Xue-Cheng Tai and Paul Tseng

Report No. 120

August 1998



Convergence Rate Analysis of an Asynchronous Space Decomposition Method for Convex Minimization*

Xue-Cheng Tai[†] Paul Tseng[‡]

August 5, 1998

Abstract

We analyze the convergence rate of an asynchronous space decomposition method for constrained convex minimization in a reflexive Banach space. This method includes as special cases parallel domain decomposition methods and multigrid methods for solving elliptic partial differential equations. In particular, the method generalizes the additive Schwarz domain decomposition methods to allow for asynchronous updates. It also generalizes the BPX multigrid method to allow for use as solvers instead of as preconditioners, possibly with asynchronous updates, and is applicable to nonlinear problems. The method is also closely related to relaxation methods for nonlinear network flow. Accordingly, we specialize our convergence rate results to the above methods. The asynchronous method is implementable in a multiprocessor system, allowing for communication and computation delays among the processors.

1 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 linear and nonlinear 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. In some systems, the activities of the processors are highly synchronized (possibly via a central processor), while in other systems (typically those with many processors), the processors may experience communication or computation delays. The latter lack of synchronization makes the analysis of the methods much more difficult. To aid in this analysis, Chazan and Miranker [14] proposed a model of asynchronous computation that allows for communication and computation delays among processors, and they showed that the Jacobi method for solving diagonally dominant system of linear equations converges under this model of asynchronous computation. Subsequently, there has been extensive study of asynchronous methods based on such a model (see [5], [6] and references therein). For these methods, convergence typically requires the algorithmic mapping to be either isotone or nonexpansive with respect to the L^∞ -norm or gradient-like. 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. One such study was done in [37] for an asynchronous gradient-projection method.

*The work of the first author was supported by the Norwegian Research Council Strategic Institute Program within Inverse Problems at RF-Rogaland Research, and by Project SEP-115837/431 at Mathematics Institute, University of Bergen. The work of the second author was supported by the National Science Foundation, Grant No. CCR-9311621.

[†]Department of Mathematics, University of Bergen, Johannes Brunsgate 12, 5007, Bergen, Norway (Xue-Cheng.Tai@mi.uib.no).

[‡]Department of Mathematics, Box 354350, University of Washington, Seattle, WA 98195, U.S.A. (tseng@math.washington.edu).

Conversion Rate Analysis of an Asynchronous Queue Based on a Method for Queue Minimization

Jun-ichi Imai, Paul Taylor

August 6, 1985

Abstract

The paper describes a method for analyzing the conversion rate of an asynchronous queue. The method is based on a queue minimization algorithm. The queue minimization algorithm is a heuristic algorithm which minimizes the number of requests in the queue. The queue minimization algorithm is applied to the queue analysis. The queue analysis is a method for analyzing the conversion rate of an asynchronous queue. The queue analysis is based on the queue minimization algorithm. The queue analysis is a method for analyzing the conversion rate of an asynchronous queue. The queue analysis is based on the queue minimization algorithm.

1. Introduction

With the advent of multiprocessor systems, queueing theory has become an important tool for analyzing the performance of such systems. In this paper, we describe a method for analyzing the conversion rate of an asynchronous queue. The queue analysis is a method for analyzing the conversion rate of an asynchronous queue. The queue analysis is based on the queue minimization algorithm. The queue analysis is a method for analyzing the conversion rate of an asynchronous queue. The queue analysis is based on the queue minimization algorithm.

This work was supported by the National Science Foundation Grant DCR-8410000. The authors would like to thank the anonymous referees for their helpful comments. The authors are also grateful to the National Science Foundation for their support of this work. The authors are also grateful to the National Science Foundation for their support of this work. The authors are also grateful to the National Science Foundation for their support of this work.

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 the form

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

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 Gâteaux-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 [4], [5], [38] and references therein) and to domain decomposition (DD) and multigrid (MG) methods for solving elliptic partial differential equations (see [7], [12], [16], [29], [34], [35], [39] 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 [7], [12], [29], [39]. 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 [1], [10], [18], [24], we especially mention the numerical tests by Frommer et al. [18] 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, \dots, K_m being suitable subspaces of V) or almost linear, this issue has been much studied for synchronous methods such as block Jacobi and block Gauss-Seidel methods (see [7], [12, §4], [29], [39, §4] and references therein) but little studied for asynchronous methods [1], [10], [24]. In the case where the equation is generally nonlinear (corresponding to K_1, \dots, K_m being suitable subspaces of V), there are some convergence studies for synchronous methods [13], [16], [28], [34], [35], and none for asynchronous methods. In the case where K_1, \dots, K_m are not all subspaces, there are some convergence studies for synchronous methods and, in particular, block Jacobi and Gauss-Seidel methods (see [22], [23], [30], [33] and references therein) 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 Gâteaux derivative F' and a norm equivalence assumption on the product of K_1, \dots, 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 (finite-dimensional) linearly constrained convex programs and, in particular, nonlinear network flow problems. This yields convergence rate results for some asynchronous network relaxation methods (see §6). The convergence result are also applied 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 §7). Our results may also apply to obstacle problems, but this would require further study. Finally, we note that alternative approaches such as Newton-type methods have also been applied to develop synchronous DD and MG methods for nonlinear partial differential equations [2], [3], [9], [20], [25], [40], [41]. However, these methods use the traditional DD and MG approach or use a special two-grid treatment.

2 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

In the case of the 2×2 matrix, the characteristic polynomial is $\lambda^2 - \text{tr}(A)\lambda + \det(A)$. The roots are $\lambda = \frac{\text{tr}(A) \pm \sqrt{\text{tr}(A)^2 - 4\det(A)}}{2}$.

$$\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$$

where $\text{tr}(A)$ is the trace of A and $\det(A)$ is the determinant of A . For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the trace is $\text{tr}(A) = a + d$ and the determinant is $\det(A) = ad - bc$. The characteristic polynomial is $\lambda^2 - (a+d)\lambda + (ad - bc) = 0$. The roots are $\lambda = \frac{a+d \pm \sqrt{(a+d)^2 - 4(ad - bc)}}{2}$.

The eigenvalues of a matrix A are the roots of its characteristic polynomial. For a 2×2 matrix, there are two eigenvalues. If the discriminant $(a+d)^2 - 4(ad - bc)$ is positive, the eigenvalues are real and distinct. If it is zero, the eigenvalues are real and repeated. If it is negative, the eigenvalues are complex conjugates.

For a 2×2 matrix A , the eigenvectors corresponding to an eigenvalue λ are found by solving $(A - \lambda I)\mathbf{v} = \mathbf{0}$, where I is the identity matrix and \mathbf{v} is the eigenvector. This is a system of linear equations with two equations and two unknowns. One equation is typically redundant, so we solve one equation for one variable in terms of the other.

For example, if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and λ is an eigenvalue, then $(A - \lambda I)\mathbf{v} = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. This gives the system $(a - \lambda)v_1 + bv_2 = 0$ and $cv_1 + (d - \lambda)v_2 = 0$. If $b \neq 0$, we can solve the first equation for $v_1 = -\frac{b}{a - \lambda}v_2$. Substituting this into the second equation gives $c(-\frac{b}{a - \lambda}v_2) + (d - \lambda)v_2 = 0$, which simplifies to $(d - \lambda) - \frac{bc}{a - \lambda} = 0$. This is the characteristic equation $(a - \lambda)(d - \lambda) - bc = 0$.

Once the eigenvalues are found, the corresponding eigenvectors can be found by substituting each eigenvalue back into the system $(A - \lambda I)\mathbf{v} = \mathbf{0}$ and solving for \mathbf{v} . For a repeated eigenvalue, the eigenvectors are not unique, and we typically choose a normalized eigenvector.

3. Problem Description and Space Decomposition

The problem is to find the eigenvalues and eigenvectors of a 2×2 matrix A . The eigenvalues are the roots of the characteristic polynomial $\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$. The eigenvectors are found by solving $(A - \lambda I)\mathbf{v} = \mathbf{0}$ for each eigenvalue λ .



minimization problem

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

where K is a nonempty closed (in the strong topology) convex set in V and $F : V \mapsto \mathfrak{R}$ is a lower semicontinuous convex Gâteaux-differentiable function. We assume F is strongly convex on K or, equivalently, its Gâteaux derivative $\lim_{t \rightarrow 0} (F(v + tw) - F(v))/t$, 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.,

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

where $\sigma > 0$. It is known that, under the above assumptions, (2) has a unique solution \bar{u} [19, p. 23].

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

$$K = \sum_{i=1}^m K_i, \quad (4)$$

for some nonempty closed convex sets K_i in V , $i = 1, \dots, m$. 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 $v_i \in K_i$, $i = 1, \dots, m$, we have $\sum_{i=1}^m v_i \in K$. Following Xu [39], 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 $(\bar{u}_1, \dots, \bar{u}_m)$ being a solution (not necessarily unique) of (1) if and only if $\bar{u}_i \in K_i$ for $i = 1, \dots, m$ 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 [39] (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 [30], [33] (also see [31], [34] 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 $v_i \in K_i$, $i = 1, \dots, m$, there exists $\bar{u}_i \in K_i$ satisfying

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

See [12, p. 95], [33], [34], [39, Lemma 7.1] 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

$$\sum_{i=1}^m \sum_{j=1}^m \langle F'(w_{ij} + u_{ij}) - F'(w_{ij}), v_i \rangle \leq C_2 \left(\sum_{j=1}^m \max_{i=1, \dots, m} \|u_{ij}\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}}, \quad (6)$$

$$\forall w_{ij} \in K, u_{ij} \in K_j^\ominus, v_i \in K_i^\ominus, i, j = 1, \dots, m,$$

where we define the set difference $K_i^\ominus = \{u - v : u, v \in K_i\} \subset V$. The above assumption generalizes those in [33], [34], [35] 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, \dots, K_m one of c colors, with the colors numbered from 1 up to c , such that sets painted the same color $k \in \{1, \dots, c\}$ are orthogonal in the sense that

$$\left\| \sum_{i \in I(k)} v_i \right\|^2 = \sum_{i \in I(k)} \|v_i\|^2, \quad \forall v_i \in K_i^\ominus, i \in I(k), \quad (7)$$

$$\left\langle F' \left(u + \sum_{i \in I(k)} v_i \right), \sum_{i \in I(k)} v_i \right\rangle \leq \sum_{i \in I(k)} \langle F'(u + v_i), v_i \rangle, \quad \forall u \in K, v_i \in K_i^\ominus, i \in I(k), \quad (8)$$

where $I(k) = \{i \in \{1, \dots, m\} : K_i \text{ is painted color } k\}$. See [12, §4.1], [35] for similar orthogonal decompositions in the case K_i is a subspace. Thus $I(1), \dots, I(c)$ are disjoint subsets of $\{1, \dots, m\}$

Let \mathcal{A} be a family of subsets of X . We say that \mathcal{A} is a σ -algebra if it satisfies the following conditions:

- (i) $X \in \mathcal{A}$.
- (ii) If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$.
- (iii) If $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$, then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

It is easy to check that the collection of all subsets of X is a σ -algebra.

Let \mathcal{A} be a σ -algebra on X . We define the σ -algebra generated by \mathcal{A} , denoted by $\sigma(\mathcal{A})$, to be the smallest σ -algebra containing \mathcal{A} . The following theorem characterizes $\sigma(\mathcal{A})$.

Theorem 1.1. Let \mathcal{A} be a family of subsets of X . Then $\sigma(\mathcal{A})$ is the intersection of all σ -algebras containing \mathcal{A} .

Proof. Let \mathcal{C} be the intersection of all σ -algebras containing \mathcal{A} . We show that \mathcal{C} is a σ -algebra containing \mathcal{A} . First, $X \in \mathcal{C}$ because $X \in \mathcal{A}$ and $\mathcal{A} \subset \mathcal{C}$. Next, if $A \in \mathcal{C}$, then $A^c \in \mathcal{C}$ because $A^c \in \mathcal{A}$ and $\mathcal{A} \subset \mathcal{C}$. Finally, if $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{C}$, then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{C}$ because $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ and $\mathcal{A} \subset \mathcal{C}$.

Conversely, let \mathcal{B} be any σ -algebra containing \mathcal{A} . Then $\mathcal{C} \subset \mathcal{B}$ because \mathcal{C} is the intersection of all such \mathcal{B} .

Therefore, $\mathcal{C} = \sigma(\mathcal{A})$.

whose union is $\{1, \dots, m\}$ 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 $\Omega \subset \mathbb{R}^d$, each set K_i may correspond to a subdomain of Ω and sets are painted different colors if their corresponding subdomains intersect (see §6, §7 for details).

Remark 1: It can be seen that condition (6) is implied by the following strengthened Cauchy-Schwarz inequality (also see [29, p. 155], [39] for the case of quadratic F and subspace K_i):

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

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

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 analysis 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 §7).

3 An Asynchronous Space Decomposition Method

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

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

(see [19, p. 23]). Let $\pi_i(u_1, \dots, u_m)$ denote this w_i . Then (π_1, \dots, π_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 $(u_1(t), \dots, u_m(t))$, $t = 0, 1, \dots$, with $(u_1(0), \dots, u_m(0)) \in K_1 \times \dots \times K_m$ given, according to the updating formula:

$$u_i(t+1) = u_i(t) + \gamma s_i(t), \quad i = 1, \dots, m, \quad (10)$$

where we define

$$s_i(t) = \begin{cases} w_i(t) - u_i(t) & \text{if } t \in T^i \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

$$w_i(t) = \pi_i(u_1(\tau_1^i(t)), \dots, u_m(\tau_m^i(t))), \quad (12)$$

and T^i is some subset of $\{0, 1, \dots\}$ 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 $(u_1(t), \dots, u_m(t)) \in K_1 \times \dots \times K_m$ for all $t = 0, 1, \dots$

We will assume that the iterates are updated in a *partially asynchronous* manner [5, Chap. 7], i.e., there exists an integer $B \geq 1$ such that

$$\{t, t+1, \dots, t+B-1\} \cap T^i \neq \emptyset \quad t = 0, 1, \dots, \forall i, \quad (13)$$

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

We say that a color $k \in \{1, \dots, c\}$ is *active* at time t if there exists an $i \in I(k)$ such that $t \in T^i$. Recall that $I(k)$ indexes those sets painted the color k . Denoting by c_t the total number of colors that are active at time t , we will also assume that

$$\gamma < \min \left\{ \frac{\sigma}{2C_2B}, \frac{1}{c_t} \right\}, \quad t = 0, 1, \dots \quad (15)$$

Notice that γ does not depend on m nor on C_1 . Although (15) may give a very conservative value of γ , this can be remedied by starting with a larger γ and decreasing γ whenever “sufficient progress” (defined in any reasonable way) is not made and (15) is not satisfied.

Let \mathcal{A} be a subalgebra of \mathcal{B} . Then \mathcal{A} is a σ -algebra if and only if \mathcal{A} is closed under countable unions and complements. In other words, \mathcal{A} is a σ -algebra if and only if \mathcal{A} is a σ -ring and \mathcal{A} contains the universal set Ω .

Let \mathcal{A} and \mathcal{B} be σ -algebras. Then $\mathcal{A} \cap \mathcal{B}$ is a σ -algebra. The σ -algebra generated by \mathcal{A} and \mathcal{B} , denoted by $\mathcal{A} \vee \mathcal{B}$, is the smallest σ -algebra containing both \mathcal{A} and \mathcal{B} .

3. An Algorithmic Approach to Probability

Let \mathcal{A} be a σ -algebra. A probability measure P on \mathcal{A} is a function $P: \mathcal{A} \rightarrow [0, 1]$ such that $P(\Omega) = 1$ and P is countably additive.

Let \mathcal{A} be a σ -algebra. A random variable X on \mathcal{A} is a measurable function $X: \Omega \rightarrow \mathbb{R}$. The distribution function of X is the function $F_X: \mathbb{R} \rightarrow [0, 1]$ defined by $F_X(x) = P(X \leq x)$.

Let X and Y be random variables. Then $X + Y$ is a random variable. The distribution function of $X + Y$ is given by $F_{X+Y}(x) = P(X + Y \leq x)$.

$$F_{X+Y}(x) = P(X + Y \leq x) = P(\omega \in \Omega : X(\omega) + Y(\omega) \leq x)$$

Let X and Y be independent random variables. Then the distribution function of $X + Y$ is given by $F_{X+Y}(x) = F_X(x) * F_Y(x)$, where $*$ denotes convolution.

$$F_{X+Y}(x) = \int_{-\infty}^{\infty} F_X(x - y) dF_Y(y)$$

Let X and Y be independent random variables. Then the characteristic function of $X + Y$ is given by $\phi_{X+Y}(t) = \phi_X(t) \phi_Y(t)$.

$$\phi_{X+Y}(t) = \int_{-\infty}^{\infty} e^{it(x+y)} dF_{X+Y}(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{itx} e^{ity} dF_X(x) dF_Y(y) = \phi_X(t) \phi_Y(t)$$

Let X and Y be independent random variables. Then the moment generating function of $X + Y$ is given by $M_{X+Y}(t) = M_X(t) M_Y(t)$.

Remark 3: The above asynchronous method models a situation in which computation is distributed over m processors with the i th 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 to its current copy of (u_1, \dots, 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.

Remark 4: The assumption that $\tau_i^i(t) = t$ can perhaps be removed through a more careful analysis, though this seems to be a reasonable assumption in practice. Intuitively, (13) says that each component u_i is updated at least once every B time units, and (14) says that the information used by processor i from processor j should not be out-of-date by more than B time units. This assumption of bounded communication and computation delay is needed for a convergence rate analysis.

4 Convergence Rate of the Asynchronous Method

In this section we prove that the iterates $(u_1(t), \dots, u_m(t))$, $t = 0, 1, \dots$, generated by the asynchronous method (10)–(15) attain linear rate of convergence, with a factor that depends on σ, C_1, C_2, c and B, γ only (see Theorem 1). While parts of our proof uses ideas from the analysis of asynchronous gradient-like methods [5, §7.5], [37], a number of new proof ideas are introduced to account for different problem assumptions and different natures of the Jacobi and Gauss-Seidel algorithmic mappings. To simplify the notation in our analysis, define

$$u(t) = \sum_{j=1}^m u_j(t), \quad z_i(t) = \sum_{j=1}^m u_j(\tau_j^i(t)), \quad (16)$$

for all i and t . If $t \in T^i$, then the definition (12) of $w_i(t)$ and the fact that $\tau_i^i(t) = t$ and F is Gâteaux-differentiable imply $w_i(t)$ satisfies the optimality condition

$$\langle F'(z_i(t) + w_i(t) - u_i(t)), v_i - w_i(t) \rangle \geq 0, \quad \forall v_i \in K_i. \quad (17)$$

Our analysis will be based on estimates given in the following two key lemmas.

Lemma 1 (Descent Estimate). *Let A_1 and A_2 be defined by*

$$A_2 = \frac{C_2^2 B^2}{\sigma}, \quad A_1 = \frac{\sigma}{4} - \gamma^2 A_2. \quad (18)$$

For $t = 0, 1, \dots$, we have

$$F(u(t+B)) \leq F(u(t)) - \gamma A_1 \sum_{j=1}^m \sum_{\tau=t}^{t+B-1} \|s_j(\tau)\|^2 + \gamma^3 A_2 \sum_{j=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2.$$

Proof. Fix any time $t \in \{0, 1, \dots\}$. Recall that c_t is the total number of colors active at time t and, without loss of generality, we assume that the first c_t colors are active. Then $s_i(t) = 0$ for all $i \in I(k)$ and $k > c_t$, so by defining

$$e_k(t) = \sum_{i \in I(k)} s_i(t)$$

and using (16), (10) and the convexity of F , we have

$$F(u(t+1)) = F\left(u(t) + \gamma \sum_{i=1}^m s_i(t)\right)$$

The first part of the paper is devoted to the study of the asymptotic behavior of the sequence of random variables X_n defined by the recurrence relation $X_{n+1} = \frac{X_n + Y_{n+1}}{2}$, where Y_n are independent random variables with a common distribution F . It is shown that X_n converges in distribution to a normal law if F is not too far from a normal law. The second part of the paper is devoted to the study of the asymptotic behavior of the sequence of random variables Z_n defined by the recurrence relation $Z_{n+1} = \frac{Z_n + Y_{n+1}}{2}$, where Y_n are independent random variables with a common distribution F . It is shown that Z_n converges in distribution to a normal law if F is not too far from a normal law.

4. Convergence Rate of the Asymptotic Method

In the present paper we study the convergence rate of the asymptotic method. Let X_n be a sequence of random variables defined by the recurrence relation $X_{n+1} = \frac{X_n + Y_{n+1}}{2}$, where Y_n are independent random variables with a common distribution F . Let F_n be the distribution function of X_n . It is shown that F_n converges to a normal law if F is not too far from a normal law. The convergence rate is studied in the present paper.

$$F_n(x) - \Phi\left(\frac{x - \mu}{\sigma}\right) = O\left(\frac{1}{n}\right)$$

where Φ is the normal distribution function, μ and σ are the mean and standard deviation of F .

The constant in the O -notation depends on the distribution F .

$$F_n(x) - \Phi\left(\frac{x - \mu}{\sigma}\right) = O\left(\frac{1}{n^2}\right)$$

if F is a normal law.

$$F_n(x) - \Phi\left(\frac{x - \mu}{\sigma}\right) = O\left(\frac{1}{n^3}\right)$$

if F is a normal law and F is not too far from a normal law.

$$F_n(x) - \Phi\left(\frac{x - \mu}{\sigma}\right) = O\left(\frac{1}{n^4}\right)$$

if F is a normal law and F is not too far from a normal law.

$$F_n(x) - \Phi\left(\frac{x - \mu}{\sigma}\right) = O\left(\frac{1}{n^5}\right)$$

if F is a normal law and F is not too far from a normal law.

$$\begin{aligned}
&= F\left(u(t) + \gamma \sum_{k=1}^{c_t} \sum_{i \in I(k)} s_i(t)\right) \\
&= F\left((1 - c_t \gamma)u(t) + \sum_{k=1}^{c_t} \gamma(u(t) + e_k(t))\right) \\
&\leq (1 - c_t \gamma)F(u(t)) + \gamma \sum_{k=1}^{c_t} F(u(t) + e_k(t)) \\
&= F(u(t)) + \gamma \sum_{k=1}^{c_t} \left(F(u(t) + e_k(t)) - F(u(t))\right). \tag{19}
\end{aligned}$$

Since $u(t) \in K$ and $u(t) + e_k(t) \in K$, the strong monotonicity of F' on K given in (3) implies

$$F(u(t)) \geq F(u(t) + e_k(t)) - \langle F'(u(t) + e_k(t)), e_k(t) \rangle + \frac{\sigma}{2} \|e_k(t)\|^2. \tag{20}$$

Define

$$\phi_j^i(t) = \sum_{k=1}^j u_k(\tau_k^i(t)) + \sum_{k=j+1}^m u_k(t), \quad j = 0, 1, \dots, m.$$

Then $\phi_0^i(t) = u(t)$ and $\phi_m^i(t) = z_i(t)$ and

$$\phi_j^i(t) - \phi_{j-1}^i(t) = u_j(\tau_j^i(t)) - u_j(t) \in K_j^\ominus, \quad j = 1, \dots, m.$$

If $t \in T^i$, then setting $v_i = u_i(t)$ in (17) and noting that $s_i(t) = w_i(t) - u_i(t)$ (see (11)), we obtain that

$$\begin{aligned}
0 &\leq -\langle F'(z_i(t) + s_i(t)), s_i(t) \rangle \\
&= -\langle F'(z_i(t) + s_i(t)) - F'(u(t) + s_i(t)), s_i(t) \rangle - \langle F'(u(t) + s_i(t)), s_i(t) \rangle \\
&= -\sum_{j=1}^m \langle F'(\phi_j^i(t) + s_i(t)) - F'(\phi_{j-1}^i(t) + s_i(t)), s_i(t) \rangle - \langle F'(u(t) + s_i(t)), s_i(t) \rangle.
\end{aligned}$$

If $t \notin T^i$, then $s_i(t) = 0$ and the above inequality holds trivially. Combining the above inequality with (7) and (8) and (20), we obtain that

$$\begin{aligned}
&\sum_{k=1}^{c_t} \left(F(u(t) + e_k(t)) - F(u(t))\right) \\
&\leq \sum_{k=1}^{c_t} \sum_{i \in I(k)} \langle F'(u(t) + s_i(t)), s_i(t) \rangle - \frac{\sigma}{2} \sum_{k=1}^{c_t} \sum_{i \in I(k)} \|s_i(t)\|^2 \\
&= \sum_{i=1}^m \langle F'(u(t) + s_i(t)), s_i(t) \rangle - \frac{\sigma}{2} \sum_{i=1}^m \|s_i(t)\|^2 \\
&\leq -\sum_{i=1}^m \sum_{j=1}^m \langle F'(\phi_j^i(t) + s_i(t)) - F'(\phi_{j-1}^i(t) + s_i(t)), s_i(t) \rangle - \frac{\sigma}{2} \sum_{i=1}^m \|s_i(t)\|^2. \tag{21}
\end{aligned}$$

Substituting (21) into (19) and using (6) yields

$$\begin{aligned}
F(u(t+1)) &\leq F(u(t)) + \gamma C_2 \left(\sum_{j=1}^m \max_{i=1, \dots, m} \|u_j(\tau_j^i(t)) - u_j(t)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(t)\|^2 \right)^{\frac{1}{2}} \\
&\quad - \gamma \frac{\sigma}{2} \sum_{i=1}^m \|s_i(t)\|^2. \tag{22}
\end{aligned}$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right)^2 =$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right) \left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right) =$$

$$\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} + \sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} =$$

$$2 \sum_{k=1}^n \frac{1}{k^2} + \frac{2}{n^2} =$$

Since $\sum_{k=1}^n \frac{1}{k^2} < 2$ and $\frac{2}{n^2} < 2$, the sum is bounded by 4. As $n \rightarrow \infty$, the sum approaches 4.

$$\lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right)^2 = 4$$

$$\sum_{k=1}^n \frac{1}{k^2} = \frac{\pi^2}{6} - \frac{1}{n} + o\left(\frac{1}{n}\right)$$

The sum $\sum_{k=1}^n \frac{1}{k^2}$ converges to $\frac{\pi^2}{6}$.

$$\lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right) = \frac{\pi^2}{6}$$

If $\epsilon > 0$, then there is a N such that for all $n > N$, $\left| \sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} - \frac{\pi^2}{6} \right| < \epsilon$.

$$\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} > \frac{\pi^2}{6} - \epsilon$$

$$\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} < \frac{\pi^2}{6} + \epsilon$$

$$\sum_{k=1}^n \frac{1}{k^2} > \frac{\pi^2}{6} - \epsilon - \frac{1}{n^2}$$

If $\epsilon > 0$, then there is a N such that for all $n > N$, $\left| \sum_{k=1}^n \frac{1}{k^2} - \frac{\pi^2}{6} \right| < \epsilon$.

$$\sum_{k=1}^n \frac{1}{k^2} > \frac{\pi^2}{6} - \epsilon$$

$$\sum_{k=1}^n \frac{1}{k^2} < \frac{\pi^2}{6} + \epsilon$$

$$\sum_{k=1}^n \frac{1}{k^2} > \frac{\pi^2}{6} - \epsilon$$

$$\sum_{k=1}^n \frac{1}{k^2} < \frac{\pi^2}{6} + \epsilon$$

Therefore, $\sum_{k=1}^n \frac{1}{k^2} \rightarrow \frac{\pi^2}{6}$ as $n \rightarrow \infty$.

$$\left(\sum_{k=1}^n \frac{1}{k^2} + \frac{1}{n^2} \right)^2 \rightarrow \left(\frac{\pi^2}{6} \right)^2 = \frac{\pi^4}{36}$$

$$\frac{\pi^4}{36}$$

Since $t - B + 1 \leq \tau_j^i(t) \leq t$ for all i and j , we also have from (10) and the triangle inequality that

$$\|u_j(\tau_j^i(t)) - u_j(t)\|^2 \leq \gamma^2 \left(\sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\| \right)^2 \leq \gamma^2 B \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2. \quad (23)$$

Combining (22) and (23) yields

$$\begin{aligned} & F(u(t+1)) \\ & \leq F(u(t)) + \gamma^2 C_2 \sqrt{B} \left(\sum_{j=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(t)\|^2 \right)^{\frac{1}{2}} - \gamma \frac{\sigma}{2} \sum_{i=1}^m \|s_i(t)\|^2 \quad (24) \\ & \leq F(u(t)) + \gamma^3 \frac{C_2^2 B}{\sigma} \sum_{j=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2 - \gamma \frac{\sigma}{4} \sum_{i=1}^m \|s_i(t)\|^2, \end{aligned}$$

where the second inequality uses the identity $ab \leq (a^2 + b^2)/2$ with a and b being the two square-root terms multiplied and divided, respectively, by $B^{1/4} \sqrt{2\gamma C_2/\sigma}$. Applying the above argument successively to $t, t+1, \dots, t+B-1$ and we obtain

$$\begin{aligned} & F(u(t+B)) - F(u(t)) \\ & \leq -\gamma \left(\frac{\sigma}{4} - \frac{\gamma^2 C_2^2 B^2}{\sigma} \right) \sum_{j=1}^m \sum_{\tau=t}^{t+B-1} \|s_j(\tau)\|^2 + \gamma^3 \frac{C_2^2 B^2}{\sigma} \sum_{j=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2. \end{aligned}$$

This proves the lemma. \blacksquare

The next key lemma estimates the optimality gap $F(u(t+B)) - F(\bar{u})$, where \bar{u} is the unique solution of (2).

Lemma 2 (*Optimality Gap Estimate*). Let A_3 and A_4 be defined by

$$A_4 = \frac{C_2 B^2}{2} + \frac{8C_1^2 C_2^2 B}{\sigma}, \quad A_3 = \frac{3C_2}{2} + \frac{6C_1^2 C_2^2}{\sigma} + A_4. \quad (25)$$

For $t = 0, 1, \dots$, we have

$$\begin{aligned} F(u(t+B)) - F(\bar{u}) & \leq (1 - \gamma)(F(u(t)) - F(\bar{u})) \\ & \quad + \gamma A_3 \sum_{j=1}^m \sum_{\tau=t}^{t+B-1} \|s_j(\tau)\|^2 + \gamma^3 A_4 \sum_{j=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_j(\tau)\|^2. \end{aligned}$$

Proof. Fix any $t \in \{0, 1, \dots\}$. For each $i \in \{1, \dots, m\}$, let t^i denote the greatest element of T^i less than $t+B$. Then, we have from (11) and (17) that

$$\langle F'(z_i(t^i) + s_i(t^i)), v_i - w_i(t^i) \rangle \geq 0, \quad \forall v_i \in K_i. \quad (26)$$

We also have from (10) and (16) that

$$\begin{aligned} u_i(t+B) & = u_i(t^i) + \gamma s_i(t^i), \\ u(t+B) & = \sum_{i=1}^m u_i(t^i + 1) = \sum_{i=1}^m u_i(t^i) + \gamma \sum_{i=1}^m s_i(t^i). \end{aligned}$$

For notational simplicity, define

$$w(t) = \sum_{i=1}^m w_i(t^i), \quad \hat{u}(t) = \sum_{i=1}^m u_i(t^i).$$

Let $\mathcal{A} = \{A_1, \dots, A_n\}$ be a family of subsets of S . Let $\mathcal{B} = \{B_1, \dots, B_m\}$ be another family of subsets of S .

$$|\mathcal{A} \cap \mathcal{B}| = \sum_{i=1}^n \sum_{j=1}^m |A_i \cap B_j|$$

$$|\mathcal{A} \cap \mathcal{B}| = \sum_{i=1}^n \sum_{j=1}^m |A_i \cap B_j|$$

$$|\mathcal{A} \cap \mathcal{B}| = \sum_{i=1}^n \sum_{j=1}^m |A_i \cap B_j|$$

When the second identity was the identity of $\mathcal{A} \cap \mathcal{B}$ with $\mathcal{A} \cap \mathcal{B}$ and the first identity was the identity of $\mathcal{A} \cap \mathcal{B}$ with $\mathcal{A} \cap \mathcal{B}$.

$$|\mathcal{A} \cap \mathcal{B}| = \sum_{i=1}^n \sum_{j=1}^m |A_i \cap B_j|$$

The proof is complete. \square

The next two lemmas establish the distributive law for set operations.

Lemma 2 (Distributive Law for Intersection)

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Proof. Let $x \in A \cap (B \cup C)$. Then $x \in A$ and $x \in B \cup C$. Since $x \in B \cup C$, we have $x \in B$ or $x \in C$. If $x \in B$, then $x \in A \cap B$. If $x \in C$, then $x \in A \cap C$. In either case, $x \in (A \cap B) \cup (A \cap C)$. Conversely, let $x \in (A \cap B) \cup (A \cap C)$. Then $x \in A \cap B$ or $x \in A \cap C$. If $x \in A \cap B$, then $x \in A$ and $x \in B$. If $x \in A \cap C$, then $x \in A$ and $x \in C$. In either case, $x \in A$ and $x \in B \cup C$, so $x \in A \cap (B \cup C)$. \square

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Lemma 3 (Distributive Law for Union)

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Proof. Let $x \in A \cup (B \cap C)$. Then $x \in A$ or $x \in B \cap C$. If $x \in A$, then $x \in A \cup B$ and $x \in A \cup C$. If $x \in B \cap C$, then $x \in B$ and $x \in C$, so $x \in A \cup B$ and $x \in A \cup C$. In either case, $x \in (A \cup B) \cap (A \cup C)$. Conversely, let $x \in (A \cup B) \cap (A \cup C)$. Then $x \in A \cup B$ and $x \in A \cup C$. If $x \in A$, then $x \in A \cup (B \cap C)$. If $x \in B$ and $x \in C$, then $x \in B \cap C$, so $x \in A \cup (B \cap C)$. \square

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

The proof is complete. \square

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$



By assumption, there exists $\bar{u}_i \in K_i$, $i = 1, \dots, m$, such that (5) holds with $v_i = w_i(t^i)$, i.e.,

$$\bar{u} = \sum_{i=1}^m \bar{u}_i \quad \text{and} \quad \left(\sum_{i=1}^m \|w_i(t^i) - \bar{u}_i\|^2 \right)^{\frac{1}{2}} \leq C_1 \|w(t) - \bar{u}\|. \quad (27)$$

Then $(\bar{u}_1, \dots, \bar{u}_m)$ is a solution of the convex program (1) and, by F being Gâteaux-differentiable, it satisfies the optimality condition

$$\sum_{i=1}^m \langle F'(\bar{u}), v_i - \bar{u}_i \rangle \geq 0, \quad \forall v_i \in K_i, \quad i = 1, \dots, m. \quad (28)$$

Defining

$$\phi_j^i(t) = \sum_{k=1}^j w_k(t^k) + \sum_{k=j+1}^m u_k(\tau_k^i(t^i)), \quad j = 0, 1, \dots, m,$$

we have that $\phi_0^i(t) = z_i(t^i)$ and $\phi_m^i(t) = w(t)$ and

$$\phi_j^i(t) - \phi_{j-1}^i(t) = w_j(t^j) - u_j(\tau_j^i(t^i)) \in K_j^\ominus, \quad j = 1, \dots, m. \quad (29)$$

Setting $v_i = \bar{u}_i$ in (26) and $v_i = w_i(t^i)$ in (28), we obtain that

$$\begin{aligned} & \left\langle F'(w(t)) - F'(\bar{u}), w(t) - \bar{u} \right\rangle \leq \left\langle F'(w(t)), w(t) - \bar{u} \right\rangle \\ & \leq \sum_{i=1}^m \left\langle F'(w(t)) - F'(z_i(t^i) + s_i(t^i)), w_i(t^i) - \bar{u}_i \right\rangle \\ & = \sum_{i=1}^m \left\langle F'(w(t)) - F'(z_i(t^i)), w_i(t^i) - \bar{u}_i \right\rangle \\ & + \sum_{i=1}^m \left\langle F'(z_i(t^i)) - F'(z_i(t^i) + s_i(t^i)), w_i(t^i) - \bar{u}_i \right\rangle \\ & = \sum_{i=1}^m \sum_{j=1}^m \left\langle F'(\phi_j^i(t)) - F'(\phi_{j-1}^i(t)), w_i(t^i) - \bar{u}_i \right\rangle \\ & + \sum_{i=1}^m \left\langle F'(z_i(t^i)) - F'(z_i(t^i) + s_i(t^i)), w_i(t^i) - \bar{u}_i \right\rangle \\ & \leq C_2 \left(\sum_{j=1}^m \max_{i=1, \dots, m} \|u_j(\tau_j^i(t^i)) - w_j(t^j)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|w_i(t^i) - \bar{u}_i\|^2 \right)^{\frac{1}{2}} \\ & + C_2 \left(\sum_{i=1}^m \|s_i(t^i)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|w_i(t^i) - \bar{u}_i\|^2 \right)^{\frac{1}{2}} \\ & \leq C_1 C_2 \left(\sum_{j=1}^m \left(4\gamma^2 B \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + 2\|s_j(t^j)\|^2 \right) \right)^{\frac{1}{2}} \|w(t) - \bar{u}\| \\ & + C_1 C_2 \left(\sum_{i=1}^m \|s_i(t^i)\|^2 \right)^{\frac{1}{2}} \|w(t) - \bar{u}\|, \end{aligned} \quad (30)$$

where the third inequality uses (6) and (29); the fourth inequality uses (27) and the fact that (see (10), (11), (13), (14)),

$$\begin{aligned} \|u_j(\tau_j^i(t^i)) - w_j(t^j)\|^2 &= \|u_j(\tau_j^i(t^i)) - u_j(t^j) - s_j(t^j)\|^2 \\ &\leq 2\|u_j(\tau_j^i(t^i)) - u_j(t^j)\|^2 + 2\|s_j(t^j)\|^2 \end{aligned}$$

1. The function $f(x)$ is defined on the interval $[0, 1]$ by the formula $f(x) = \sqrt{x}$.

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

2. The function $f(x)$ is defined on the interval $[0, 1]$ by the formula $f(x) = \sqrt{x}$.

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

3. The function $f(x)$ is defined on the interval $[0, 1]$ by the formula $f(x) = \sqrt{x}$.

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$

4. The function $f(x)$ is defined on the interval $[0, 1]$ by the formula $f(x) = \sqrt{x}$.

$$f(x) = \sqrt{x} \Rightarrow f'(x) = \frac{1}{2\sqrt{x}}$$



$$\begin{aligned}
&\leq 2\gamma^2 \left(\sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\| \right)^2 + 2\|s_j(t^j)\|^2 \\
&\leq 4\gamma^2 B \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + 2\|s_j(t^j)\|^2.
\end{aligned}$$

Also, the strong monotonicity (3) of F' on K implies

$$\langle F'(w(t)) - F'(\bar{u}), w(t) - \bar{u} \rangle \geq \sigma \|w(t) - \bar{u}\|^2,$$

which together with (30) yields

$$\begin{aligned}
\|w(t) - \bar{u}\| &\leq \frac{C_1 C_2}{\sigma} \left(\sum_{j=1}^m \left(4\gamma^2 B \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + 2\|s_j(t^j)\|^2 \right) \right)^{\frac{1}{2}} \\
&\quad + \frac{C_1 C_2}{\sigma} \left(\sum_{i=1}^m \|s_i(t^i)\|^2 \right)^{\frac{1}{2}}. \tag{31}
\end{aligned}$$

Next, since $F'(w(t))$ is a subgradient of F at $w(t)$ [17, p. 23], we have

$$F(w(t)) - F(\bar{u}) \leq \langle F'(w(t)), w(t) - \bar{u} \rangle,$$

so putting $v_i = \bar{u}_i$ in (26) and adding it to the above inequality yields

$$\begin{aligned}
&F(w(t)) - F(\bar{u}) \\
&\leq \sum_{i=1}^m \langle F'(w(t)) - F'(z_i(t^i) + s_i(t^i)), w_i(t^i) - \bar{u}_i \rangle \\
&\leq \frac{C_1^2 C_2^2}{\sigma} \left(\left(\sum_{j=1}^m \left(4\gamma^2 B \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + 2\|s_j(t^j)\|^2 \right) \right)^{\frac{1}{2}} + \left(\sum_{i=1}^m \|s_i(t^i)\|^2 \right)^{\frac{1}{2}} \right)^2 \\
&\leq \frac{2C_1^2 C_2^2}{\sigma} \left(4\gamma^2 B \sum_{j=1}^m \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + 3 \sum_{i=1}^m \|s_i(t^i)\|^2 \right), \tag{32}
\end{aligned}$$

where the second inequality uses (30) and (31) and the last inequality follows from the identity $(a+b)^2 \leq 2(a^2+b^2)$.

Next we estimate $F(\hat{u}(t)) - F(u(t))$. Let $\bar{t} = \max_{i=1, \dots, m} t^i$ and, for each $i \in \{1, \dots, m\}$ and $\tau \in \{t, \dots, \bar{t}\}$, define

$$\bar{u}_i(\tau) = u_i(\min\{\tau, t^i\}), \quad \bar{u}(\tau) = \sum_{i=1}^m \bar{u}_i(\tau). \tag{33}$$

Then, for each $i \in \{1, \dots, m\}$ and $\tau \in \{t, \dots, \bar{t} - 1\}$, either $\bar{u}_i(\tau + 1) = \bar{u}_i(\tau)$ so that

$$\langle F'(z_i(\tau) + s_i(\tau)), \bar{u}_i(\tau) - \bar{u}_i(\tau + 1) \rangle = 0$$

or $\bar{u}_i(\tau + 1) \neq \bar{u}_i(\tau)$ so that $\tau \in T^i$ and $\tau < t^i$, implying by (11) and (17) that

$$\langle F'(z_i(\tau) + s_i(\tau)), u_i(\tau) - w_i(\tau) \rangle \geq 0$$

and hence, by (33), that

$$\begin{aligned}
\langle F'(z_i(\tau) + s_i(\tau)), \bar{u}_i(\tau) - \bar{u}_i(\tau + 1) \rangle &= \langle F'(z_i(\tau) + s_i(\tau)), u_i(\tau) - u_i(\tau + 1) \rangle \\
&= \gamma \langle F'(z_i(\tau) + s_i(\tau)), u_i(\tau) - w_i(\tau) \rangle \geq 0.
\end{aligned}$$

Using this and defining

$$\phi_j^i(\tau) = \sum_{k=1}^j \bar{u}_k(\tau + 1) + \sum_{k=j+1}^m u_k(\tau_k^i(\tau)), \quad j = 0, 1, \dots, m,$$

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

$$f(x) = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n$$

Let the series expansion of $f(x)$ be $\sum_{n=0}^{\infty} c_n x^n$.

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n$$

which together with (1) yields

$$\left(\sum_{n=0}^{\infty} c_n x^n \right) = \left(\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n \right)$$

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

Now, since $f(x)$ is a rational function of x , it will be a sum of two

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

in partial fractions in (2) and substituting in the above identity yields

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

$$\left(\sum_{n=0}^{\infty} c_n x^n \right) = \left(\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n \right)$$

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

where the series expansion of $f(x)$ and (2) and (3) and we have proved that the identity

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

holds for all x in the interval $|x| < R$, and for each $n = 0, 1, 2, \dots$ and

$$c_n = a_n + b_n$$

$$\sum_{n=0}^{\infty} c_n x^n = \sum_{n=0}^{\infty} (a_n + b_n) x^n$$

Thus, for each $n = 0, 1, 2, \dots$, we have $c_n = a_n + b_n$.

$$c_n = a_n + b_n$$

where $a_n = \frac{1}{2\pi i} \int_{\gamma} f(z) z^{-n-1} dz$ and $b_n = \frac{1}{2\pi i} \int_{\gamma} g(z) z^{-n-1} dz$.

$$c_n = a_n + b_n$$

and hence, by (2), that

$$\frac{1}{2\pi i} \int_{\gamma} f(z) z^{-n-1} dz = \frac{1}{2\pi i} \int_{\gamma} (f(z) + g(z)) z^{-n-1} dz$$

$$= \frac{1}{2\pi i} \int_{\gamma} f(z) z^{-n-1} dz + \frac{1}{2\pi i} \int_{\gamma} g(z) z^{-n-1} dz$$

Using the residue theorem

$$c_n = \sum_{k=0}^n a_k + \sum_{k=0}^n b_k$$

we obtain that

$$\begin{aligned}
& F(\tilde{u}(\tau+1)) - F(\tilde{u}(\tau)) \\
\leq & -\langle F'(\tilde{u}(\tau+1)), \tilde{u}(\tau) - \tilde{u}(\tau+1) \rangle \\
\leq & \sum_{i=1}^m \langle F'(z_i(\tau) + s_i(\tau)) - F'(\tilde{u}(\tau+1)), \tilde{u}_i(\tau) - \tilde{u}_i(\tau+1) \rangle \\
= & \sum_{i=1}^m \langle F'(z_i(\tau)) - F'(\tilde{u}(\tau+1)), \tilde{u}_i(\tau) - \tilde{u}_i(\tau+1) \rangle \\
+ & \sum_{i=1}^m \langle F'(z_i(\tau) + s_i(\tau)) - F'(z_i(\tau)), \tilde{u}_i(\tau) - \tilde{u}_i(\tau+1) \rangle \\
= & \sum_{i=1}^m \sum_{j=1}^m \langle F'(\phi_{j-1}^i(\tau)) - F'(\phi_j^i(\tau)), \tilde{u}_i(\tau) - \tilde{u}_i(\tau+1) \rangle \\
+ & \sum_{i=1}^m \langle F'(z_i(\tau) + s_i(\tau)) - F'(z_i(\tau)), \tilde{u}_i(\tau) - \tilde{u}_i(\tau+1) \rangle \\
\leq & C_2 \left(\sum_{j=1}^m \max_{i=1, \dots, m} \|\phi_{j-1}^i(\tau) - \phi_j^i(\tau)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|\tilde{u}_i(\tau) - \tilde{u}_i(\tau+1)\|^2 \right)^{\frac{1}{2}} \\
+ & C_2 \left(\max_{i=1, \dots, m} \|s_i(\tau)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|\tilde{u}_i(\tau) - \tilde{u}_i(\tau+1)\|^2 \right)^{\frac{1}{2}} \tag{34} \\
\leq & C_2 \gamma \left(\sum_{j=1}^m \left(\max_{i=1, \dots, m} \|\tilde{u}_j(\tau+1) - u_j(\tau_j^i(\tau))\|^2 \right) \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(\tau)\|^2 \right)^{\frac{1}{2}} \\
+ & C_2 \gamma \left(\sum_{i=1}^m \|s_i(\tau)\|^2 \right) \\
\leq & \gamma C_2 \left(\gamma^2 B \sum_{j=1}^m \sum_{\nu=\tau-B+1}^{\tau+1} \|s_j(\nu)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(\tau)\|^2 \right)^{\frac{1}{2}} + C_2 \gamma \left(\sum_{i=1}^m \|s_i(\tau)\|^2 \right) \\
\leq & \gamma^3 \frac{C_2 B}{2} \sum_{j=1}^m \sum_{\nu=\tau-B+1}^{\tau+1} \|s_j(\nu)\|^2 + \gamma \frac{3C_2}{2} \sum_{i=1}^m \|s_i(\tau)\|^2,
\end{aligned}$$

where the first inequality uses the subgradient property of $F'(\tilde{u}(\tau+1))$ [17, p. 23]; the third inequality uses (6); the fourth and fifth inequalities use (33) and (10) and an inequality analogous to (23); the last inequality uses the identity $ab \leq (a^2 + b^2)/2$ with a and b being the two square-root terms. Summing the above inequality over $\tau = t, t+1, \dots, \bar{t}-1$ and observing that $\tilde{u}(\bar{t}) = \hat{u}(t)$ and $\tilde{u}(t) = u(t)$, we then have

$$\begin{aligned}
F(\hat{u}(t)) - F(u(t)) & \leq \gamma^3 \frac{C_2 B}{2} \sum_{j=1}^m \sum_{\tau=t}^{\bar{t}-1} \sum_{\nu=\tau-B+1}^{\tau+1} \|s_j(\nu)\|^2 + \gamma \frac{3C_2}{2} \sum_{i=1}^m \sum_{\tau=t}^{\bar{t}-1} \|s_i(\tau)\|^2 \\
& \leq \gamma^3 \frac{C_2 B^2}{2} \sum_{j=1}^m \sum_{\tau=t-B+1}^{t+B-1} \|s_j(\tau)\|^2 + \gamma \frac{3C_2}{2} \sum_{i=1}^m \sum_{\tau=t}^{t+B-1} \|s_i(\tau)\|^2. \tag{35}
\end{aligned}$$

Finally, using the convexity of F and $\gamma \in [0, 1]$, we see from (11) and (32) and (35) that

$$\begin{aligned}
& F(u(t+B)) - F(\bar{u}) \\
= & F \left(\sum_{i=1}^m u_i(t+B) \right) - F(\bar{u})
\end{aligned}$$

$$\begin{aligned}
 & \frac{1}{(1-x)^2} = \sum_{n=0}^{\infty} (n+1)x^n \\
 & \frac{1}{(1-x)^3} = \sum_{n=0}^{\infty} \binom{n+2}{2} x^n \\
 & \frac{1}{(1-x)^4} = \sum_{n=0}^{\infty} \binom{n+3}{3} x^n \\
 & \vdots \\
 & \frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{k} x^n
 \end{aligned}$$

where the first equation was the substitution property of $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ and the second was the binomial theorem. The induction step is similar to the first step. For the induction step, we assume that the formula holds for k and we show it holds for $k+1$. We start with the binomial theorem:

$$\begin{aligned}
 \frac{1}{(1-x)^{k+1}} &= \frac{1}{(1-x)^k} \cdot \frac{1}{1-x} \\
 &= \left(\sum_{n=0}^{\infty} \binom{n+k-1}{k-1} x^n \right) \left(\sum_{m=0}^{\infty} x^m \right) \\
 &= \sum_{n=0}^{\infty} \left(\sum_{m=0}^n \binom{n+k-1}{k-1} \right) x^n
 \end{aligned}$$

Using the identity $\sum_{m=0}^n \binom{n+k-1}{k-1} = \binom{n+k}{k}$, we see that the coefficient of x^n in the product is $\binom{n+k}{k}$, which is exactly the coefficient of x^n in the series $\sum_{n=0}^{\infty} \binom{n+k}{k} x^n$. This completes the induction step.



$$\begin{aligned}
&= F\left(\sum_{i=1}^m (u_i(t^i) + \gamma(w_i(t^i) - u_i(t^i)))\right) - F(\bar{u}) \\
&= F((1-\gamma)\hat{u}(t) + \gamma w(t)) - F(\bar{u}) \\
&\leq (1-\gamma)F(\hat{u}(t)) + \gamma F(w(t)) - F(\bar{u}) \\
&= (1-\gamma)(F(\hat{u}(t)) - F(\bar{u})) + \gamma(F(w(t)) - F(\bar{u})) \\
&\leq (1-\gamma)(F(u(t)) - F(\bar{u})) + \gamma^3 \frac{C_2 B^2}{2} \sum_{j=1}^m \sum_{\tau=t-B+1}^{t+B-1} \|s_j(\tau)\|^2 + \gamma \frac{3C_2}{2} \sum_{i=1}^m \sum_{\tau=t}^{t+B-1} \|s_i(\tau)\|^2 \\
&+ \gamma^3 \frac{8C_1^2 C_2^2 B}{\sigma} \sum_{j=1}^m \sum_{\tau=t-B+1}^{t+B-2} \|s_j(\tau)\|^2 + \gamma \frac{6C_1^2 C_2^2}{\sigma} \sum_{i=1}^m \|s_i(t^i)\|^2.
\end{aligned}$$

Using $\gamma \leq 1$ then proves the lemma. \blacksquare

We will now use Lemmas 1 and 2 to prove our convergence rate result. To simplify the notations, define

$$a_k = F(u(kB)) - F(\bar{u}), \quad b_k = \sum_{j=1}^m \sum_{\tau=kB-B}^{kB-1} \|s_j(\tau)\|^2, \quad k = 1, 2, \dots$$

By Lemmas 1 and 2, we have

$$a_k \leq a_{k-1} - \gamma A_1 b_k + \gamma^3 A_2 b_{k-1}, \quad (36)$$

$$a_k \leq (1-\gamma)a_{k-1} + \gamma A_3 b_k + \gamma^3 A_4 b_{k-1}, \quad (37)$$

where A_1, A_2, A_3, A_4 are given by (18) and (25). By (15), we have $A_1 > 0$. Choose γ sufficiently small so that

$$\varrho = \max \left\{ \left(1 + \frac{A_1}{A_3}\right)^{-1} \left(1 + (1-\gamma)\frac{A_1}{A_3} + \gamma^{3/2} \left(A_2 + \frac{A_1 A_4}{A_3}\right)\right), A_1^{-1}(\gamma^{1/2} + \gamma^2 A_2) \right\} < 1. \quad (38)$$

Also, define $a = \max\{a_1, \gamma^{3/2} b_1\}/\varrho$. We claim that

$$\max\{a_n, \gamma^{3/2} b_n\} \leq a \varrho^n \quad (39)$$

for $n = 1, 2, \dots$. We prove this by induction on n . Clearly (39) holds for $n = 1$ by our definition of a . Suppose (39) holds for $n = k - 1$, where $k > 1$. Multiplying (37) by A_1/A_3 and adding it to (36) gives

$$\left(1 + \frac{A_1}{A_3}\right) a_k \leq \left(1 + (1-\gamma)\frac{A_1}{A_3}\right) a_{k-1} + \gamma^{3/2} \left(A_2 + \frac{A_1 A_4}{A_3}\right) (\gamma^{3/2} b_{k-1}),$$

which together with the inductive hypothesis $\max\{a_{k-1}, \gamma^{3/2} b_{k-1}\} \leq a \varrho^{k-1}$ and (38) yields

$$a_k \leq \left(1 + \frac{A_1}{A_3}\right)^{-1} \left(1 + (1-\gamma)\frac{A_1}{A_3} + \gamma^{3/2} \left(A_2 + \frac{A_1 A_4}{A_3}\right)\right) a \varrho^{k-1} \leq a \varrho^k.$$

Similarly, (36) and $a_k \geq 0$ give

$$\gamma^{3/2} A_1 b_k \leq \gamma^{1/2} a_{k-1} + \gamma^2 A_2 (\gamma^{3/2} b_{k-1}),$$

which together with $\max\{a_{k-1}, \gamma^{3/2} b_{k-1}\} \leq a \varrho^{k-1}$ and (38) yields

$$\gamma^{3/2} b_k \leq A_1^{-1}(\gamma^{1/2} + \gamma^2 A_2) a \varrho^{k-1} \leq a \varrho^k.$$

This shows that (39) holds for $n = k$, completing our induction proof.

Thus, we have shown linear rate of convergence (in the root sense) for both a_n and b_n , with a factor of ϱ . The latter implies $u_i(t)$, $t = 0, 1, \dots$, is a Cauchy sequence for each i and hence it converges strongly. This is summarized in the theorem below.

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k (k+1)^2$$

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k (k^2 + 2k + 1)$$

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k k^2 + 2 \sum_{k=0}^n \binom{n}{k} (-1)^k k + \sum_{k=0}^n \binom{n}{k} (-1)^k$$

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k k^2 + 2 \sum_{k=0}^n \binom{n}{k} (-1)^k k + \sum_{k=0}^n \binom{n}{k} (-1)^k$$

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k k^2 + 2 \sum_{k=0}^n \binom{n}{k} (-1)^k k + \sum_{k=0}^n \binom{n}{k} (-1)^k$$

$$= \sum_{k=0}^n \binom{n}{k} (-1)^k k^2 + 2 \sum_{k=0}^n \binom{n}{k} (-1)^k k + \sum_{k=0}^n \binom{n}{k} (-1)^k$$

Using the binomial theorem, we have

We will use the binomial theorem to find the sum of the series. To do this, we consider the series

$$f(x) = \sum_{k=0}^n \binom{n}{k} (-1)^k (k+1)^2 x^k$$

By Lemma 1 and 2, we have

$$f(x) = \sum_{k=0}^n \binom{n}{k} (-1)^k (k+1)^2 x^k$$

where $f(x)$ is the sum of the series. We will use the binomial theorem to find the sum of the series.

$$f(x) = \sum_{k=0}^n \binom{n}{k} (-1)^k (k+1)^2 x^k$$

The series is a binomial series. We can use the binomial theorem to find the sum of the series.

$$f(x) = \sum_{k=0}^n \binom{n}{k} (-1)^k (k+1)^2 x^k$$

Let $n=1$. The series is a binomial series. We can use the binomial theorem to find the sum of the series.

$$f(x) = \sum_{k=0}^1 \binom{1}{k} (-1)^k (k+1)^2 x^k$$

which together with Lemma 1 and 2, we have

$$f(x) = \sum_{k=0}^1 \binom{1}{k} (-1)^k (k+1)^2 x^k$$

Lemma 1 and 2, we have

$$f(x) = \sum_{k=0}^1 \binom{1}{k} (-1)^k (k+1)^2 x^k$$

which together with Lemma 1 and 2, we have

$$f(x) = \sum_{k=0}^1 \binom{1}{k} (-1)^k (k+1)^2 x^k$$

This shows that (1) holds for $n=1$. We will use induction to show that (1) holds for all n .

Let n be a positive integer. We will use induction to show that (1) holds for all n . The base case is $n=1$, which we have already shown. Assume that (1) holds for $n-1$. We will show that (1) holds for n .

Theorem 1 Consider the minimization problem (2) and the space decomposition (4) of §2 (see (3), (5), (6), (7)). Let $(u_1(t), \dots, u_m(t))$, $t = 0, 1, \dots$, be generated by the asynchronous space decomposition method of §3 (see (10)–(12) and (13), (14)) with stepsize γ satisfying (15), (38). Then, there exist $a > 0$ and $\rho \in (0, 1)$, depending on σ, C_1, C_2 and B, γ only, such that

$$F(u(nB)) - F(\bar{u}) \leq a\rho^n, \quad n = 1, 2, \dots,$$

where $u(t)$ is given by (16) and \bar{u} denotes the unique solution of (2). Moreover, $u(t)$ converges strongly to \bar{u} and, for each $i \in \{1, \dots, m\}$, $u_i(t)$ converges strongly as $t \rightarrow \infty$.

5 Convergence Rate of Block Jacobi and Gauss-Seidel Methods

It is readily seen that the following block Jacobi method is a special case of the asynchronous space decomposition method (10)–(12) with $T^i = \{0, 1, \dots\}$ and $\tau_j^i(t) = t$ for all i, j, t (so $B = 1$ and $c_t = c$). Thus, Theorem 1 can be applied to establish its linear convergence and obtain estimate of the factor ρ under the assumptions of §2. Moreover, by observing that in this case the left-hand side of (23) is zero so that Lemma 1 holds with $A_2 = 0$, the stepsize restriction (15) can be relaxed to $\gamma \leq 1/c_t$.

Algorithm 1

Step 1. Choose initial values $u_i(0) \in K_i$, $i = 1, \dots, m$, and stepsize $\gamma = 1/c$, where c is defined as in §2.

Step 2. For each $t = 0, 1, \dots$, find $w_i(t) \in K_i$ in parallel for $i = 1, \dots, m$ that satisfies

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

Step 3. Set

$$u_i(t+1) = u_i(t) + \gamma(w_i(t) - u_i(t)),$$

and go to the next iteration.

The following block Gauss-Seidel method is also a special case of the asynchronous space decomposition method (10)–(12) with $\gamma = 1$, $T^i = \{i - 1 + km\}_{k=0,1,\dots}$ and $\tau_j^i(t) = t$ for all i, j, t (so $B = m$ and $c_t = 1$). Here Theorem 1 cannot be directly applied due to $\gamma = 1$ possibly violating (15). However, by observing that in this case the left-hand side of (23) is again zero so that Lemma 1 holds with $A_2 = 0$, the proof of the theorem can be easily modified to establish linear convergence of this method under the assumptions of §2, with factor ρ depending on m, σ, C_1, C_2 only. Moreover, by grouping sets of the same color into one set, we can ensure that $m = c$, where c is defined as in §2.

Algorithm 2

Step 1. Choose initial values $u_i(0) \in K_i$, $i = 1, \dots, m$.

Step 2. For each $t = 0, 1, \dots$, find $u_i(t+1) \in K_i$ sequentially for $i = 1, \dots, m$ that satisfies

$$\begin{aligned} & F\left(\sum_{j < i} u_j(t+1) + u_i(t+1) + \sum_{j > i} u_j(t)\right) \\ & \leq F\left(\sum_{j < i} u_j(t+1) + v_i + \sum_{j > i} u_j(t)\right), \quad \forall v_i \in K_i. \end{aligned}$$

Step 3. Go to the next iteration.

Algorithm 1 (continued) ...

$$x_{k+1} = \frac{1}{\alpha} \left(\sum_{i=1}^n x_i + \alpha \sum_{i=1}^n x_i \right)$$

where $\alpha = \frac{1}{1 + \sqrt{1 + 4 \sum_{i=1}^n x_i}}$ and $\sum_{i=1}^n x_i$ is the sum of the components of x_k .

5. Convergence Rate of Block Jacobi and Gauss-Seidel Methods

In this section we analyze the convergence rate of the block Jacobi and Gauss-Seidel methods. We consider the block Jacobi method with m blocks of size $p \times p$ and $n = mp$. The convergence rate is determined by the spectral radius of the iteration matrix B . For the block Jacobi method, the iteration matrix is given by $B = I - A^{-1}A$, where A is the coefficient matrix. The convergence rate is $\rho(B)$.

Algorithm 1

Step 1: Initialize $x_0 = 0$, $x_1 = 0, \dots, x_m = 0$ and repeat until convergence.

Step 2: For $k = 1, 2, \dots$ do until convergence:

$$x_{k+1} = \frac{1}{\alpha} \left(\sum_{i=1}^m x_i + \alpha \sum_{i=1}^m x_i \right)$$

end do

$$\alpha = \frac{1}{1 + \sqrt{1 + 4 \sum_{i=1}^m x_i}}$$

where $\sum_{i=1}^m x_i$ is the sum of the components of x_k .

The following theorem provides an upper bound for the convergence rate of the block Jacobi method. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of the iteration matrix B . Then the convergence rate is bounded by $\max_{i=2, \dots, n} |\lambda_i|$. For the Gauss-Seidel method, the convergence rate is bounded by $\max_{i=2, \dots, n} |\lambda_i|$.

Algorithm 2

Step 1: Initialize $x_0 = 0$, $x_1 = 0, \dots, x_m = 0$ and repeat until convergence.

Step 2: For $k = 1, 2, \dots$ do until convergence:

$$x_{k+1} = \frac{1}{\alpha} \left(\sum_{i=1}^m x_i + \alpha \sum_{i=1}^m x_i \right)$$

$$x_{k+1} = \frac{1}{\alpha} \left(\sum_{i=1}^m x_i + \alpha \sum_{i=1}^m x_i \right)$$

end do

The above two methods for solving (2) were studied in [30] (also see [31], [32], [33]), where convergence of the methods was proved under weaker assumptions. However, no rate of convergence result was given. In [34], linear rate of convergence for the above two methods was proved for the unconstrained case of $K = V$. In the finite-dimensional case of $V = \mathbb{R}^n$, linear rate of convergence for the Gauss-Seidel method can also be inferred from the results in [22], [23] and references therein, but our estimate of the convergence factor is new.

In [34], the minimization subproblem at each iteration is solved inexactly. We can do likewise in the constrained case. In particular, the proof of Theorem 1 (see (21) and (26)) suggests that the exact minimization condition (17) can be relaxed to the following inexact minimization condition

$$\langle F'(z_i(t) + w_i(t) - u_i(t)), v_i - w_i(t) \rangle \geq -\frac{\sigma_0}{2} \|w_i(t) - u_i(t)\|^2, \quad \forall v_i \in K_i,$$

with $0 < \sigma_0 < \sigma$. However, σ would need to be known explicitly and both γ and ϱ would depend on σ_0 .

6 Applications to Convex Programming

In this section we consider the Euclidean space $V = V' = \mathbb{R}^n$, which is the space of n -dimensional real column vectors with duality pairing $\langle f, x \rangle = f^T x$ and norm $\|x\| = \sqrt{x^T x}$, where x^T denotes transpose of x . We will discuss choices of the space decomposition (4) and the corresponding estimates for C_1 , C_2 , c in (5), (6), (7). In the case of nonlinear network flow, we will also relate our asynchronous method to those studied in [5, §7.2.3], [38].

6.1 Primal Applications

Consider the problem (2), where $F : \mathbb{R}^n \rightarrow \mathbb{R}$ is a differentiable convex function and K is a nonempty polyhedral set in \mathbb{R}^n . Then F is continuous [26, p. 82] and continuously differentiable [26, p. 246]. 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 result of Hoffman on the Lipschitzian behavior of solutions of a linear system with respect to the right-hand side (see [11]) implies that, for any $v_i \in K_i$, $i = 1, \dots, m$, there exists $\bar{u}_i \in K_i$ satisfying (5), where the constant C_1 depends on m and certain condition numbers for K_i , $i = 1, \dots, m$. In cases where each K_i has a simple structure, such as the Cartesian product of closed intervals, C_1 may be estimated explicitly. For a coloring of the sets, if K_i and K_j are not orthogonal, i.e., $(v_i)^T v_j \neq 0$ for some $v_i \in K_i, v_j \in K_j$, then we paint them different colors. Let \hat{c} be the maximum number of sets K_j that are not orthogonal to an arbitrary set K_i . Then an analysis similar to that used in §7.1.3 shows that (6) holds with $C_2 = L\hat{c}$, where L is the Lipschitz constant for F' .

6.2 Dual Applications

Consider the linearly constrained convex program

$$\text{minimize } G(x) \quad \text{subject to } Ax = b, \quad (40)$$

where $G : \mathbb{R}^n \rightarrow \mathbb{R}$ is a strictly convex differentiable function, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$ has nonzero rows. We assume there exists $\bar{x} \in \mathbb{R}^n$ satisfying $A\bar{x} = b$. By attaching Lagrange multipliers $\lambda \in \mathbb{R}^m$ to the equations $Ax = b$ in (40), we obtain the Lagrangian dual problem:

$$\min_{\lambda \in \mathbb{R}^m} G^*(A^T \lambda) - b^T \lambda, \quad (41)$$

where G^* is the convex conjugate (also called Legendre-Fenchel transform) of G defined by (see [19], [26])

$$G^*(u) = \sup_{x \in \mathbb{R}^n} \{u^T x - G(x)\}.$$

we show two methods for finding the minimum of the function $f(x)$ on the interval $[a, b]$. The first method is the golden section search, which is based on the fact that the function $f(x)$ is unimodal. The second method is the Fibonacci search, which is based on the fact that the function $f(x)$ is unimodal and the interval $[a, b]$ is divided into segments of length F_n , where F_n is the n -th Fibonacci number. The golden section search is more efficient than the Fibonacci search because it requires fewer function evaluations. The Fibonacci search is more efficient than the golden section search because it requires fewer function evaluations.

$$f(x) = x^2 - 2x + 1, \quad x \in [0, 1]$$

with $f(x)$ as the objective function and x as the design variable.

6 Applications to Convex Programming

In this section we consider the problem of minimizing a convex function $f(x)$ over a convex set S . We will discuss the gradient method, the Newton method, and the interior point method. The gradient method is based on the fact that the gradient of a convex function points in the direction of steepest ascent. The Newton method is based on the fact that the Hessian of a convex function is positive definite. The interior point method is based on the fact that the level sets of a convex function are convex.

6.1 Gradient Method

The gradient method is a first-order method for minimizing a convex function $f(x)$ over a convex set S . It is based on the fact that the gradient of a convex function points in the direction of steepest ascent. The method starts at a point $x_0 \in S$ and iteratively moves in the direction of the negative gradient of $f(x)$ at x_0 . The method is simple to implement and requires only first-order derivatives of $f(x)$. However, it is slow to converge and may require many iterations to reach the minimum. The method is most effective when the function $f(x)$ is smooth and the set S is compact.

6.2 Dual Method

The dual method is a second-order method for minimizing a convex function $f(x)$ over a convex set S . It is based on the fact that the Hessian of a convex function is positive definite. The method starts at a point $x_0 \in S$ and iteratively moves in the direction of the negative Hessian of $f(x)$ at x_0 . The method is more efficient than the gradient method because it uses second-order derivatives of $f(x)$. However, it is more complex to implement and requires more function evaluations. The method is most effective when the function $f(x)$ is smooth and the set S is compact.

The convex programs (40) and (41) are dual in the sense that one has a solution if and only if the other does and these solutions satisfy $G'(x) = A^T \lambda$ [26, Cor. 28.3.1 and 28.4.1]. Using $b = A\bar{x}$, we can rewrite the dual problem (41) in the form of (2) with

$$F(u) = G^*(u) - \bar{x}^T u, \quad K = \{u \in \mathfrak{R}^n : u = A^T \lambda \text{ for some } \lambda \in \mathfrak{R}^m\}. \quad (42)$$

We assume that $(G^*)'$ is strongly monotone and Lipschitz continuous on \mathfrak{R}^n , so that F satisfies (3) for some $\sigma > 0$. If G is twice differentiable, this assumption essentially amounts to G'' having bounded eigenvalues and the Hessian $(G'')^{-1}$ having bounded entries on \mathfrak{R}^n . Let \bar{u} denote the unique solution of (1) and let A_i denote the i th row of A .

We can decompose K in the form (4) with subspaces

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

First we show that, for any $v_i \in K_i$, $i = 1, \dots, m$, there exists $\bar{u}_i \in K_i$ satisfying (5), where

$$C_1 = \|D^{-1} B (B^T D^{-2} B)^{-1}\|, \quad (43)$$

with D being the diagonal matrix with diagonal entries $\|A_i^T\|$, $i = 1, \dots, m$, and B being any submatrix of A comprising linearly independent columns of A spanning the column space of A . To see this, notice that $\bar{u} = A^T \bar{\lambda}$ for some $\bar{\lambda} \in \mathfrak{R}^m$ and $v_i = A_i^T \mu_i$ for some $\mu_i \in \mathfrak{R}$. Moreover, $u_i \in K_i$, $i = 1, \dots, m$, satisfy $\sum_{i=1}^m u_i = \bar{u}$ if and only if $u_i = A_i^T \lambda_i$ and $A^T \lambda = A^T \bar{\lambda}$ for some $\lambda = (\lambda_i)_{i=1}^m$. Thus, minimizing $\sum_{i=1}^m \|u_i - v_i\|^2$ subject to $u_i \in K_i$ and $\sum_{i=1}^m u_i = \bar{u}$ is equivalent to minimizing

$$\sum_{i=1}^m \|A_i^T \lambda_i - A_i^T \mu_i\|^2 = \sum_{i=1}^m \|A_i^T\|^2 |\lambda_i - \mu_i|^2 = \|D(\lambda - \mu)\|^2$$

subject to $A^T \lambda = A^T \bar{\lambda}$, where $\mu = (\mu_i)_{i=1}^m$. This in turn is equivalent to minimizing $\|D(\lambda - \mu)\|^2$ subject to $B^T \lambda = B^T \bar{\lambda}$, whose solution is $\lambda = \mu + D^{-2} B (B^T D^{-2} B)^{-1} B^T (\bar{\lambda} - \mu)$. Then

$$\begin{aligned} \|D(\lambda - \mu)\| &= \|D^{-1} B (B^T D^{-2} B)^{-1} B^T (\bar{\lambda} - \mu)\| \\ &\leq C_1 \|B^T (\bar{\lambda} - \mu)\| \leq C_1 \|A^T (\bar{\lambda} - \mu)\| = C_1 \left\| \bar{u} - \sum_{i=1}^m v_i \right\|. \end{aligned}$$

The formula for C_1 (43) simplifies if A has full row rank, in which case B is square and invertible. If A does not have full row rank, we could remove the redundant rows, but our experience with network flow problems suggests that this removal can slow the convergence of Gauss-Seidel methods on the problem [38]. 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, \dots, K_m as discussed in §6.1 and show 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 . If we replace the equation $Ax = b$ in (40) by an inequality $Ax \leq b$, we would have an additional constraint of $\lambda \leq 0$ in (41), so K would not be a subspace and the estimate (43) would need to be modified accordingly.

In the case of a nonlinear network flow problem [27], where A is the node-arc incidence matrix for a connected digraph with m nodes and n arcs, i.e., every column of A has one 1 and one -1 in two of its rows, and a 0 in the remaining rows, we can estimate C_1 explicitly in terms of m and n as follows: For any $v_i \in K_i$, $i = 1, \dots, m$, we have $\bar{u} = A^T \bar{\lambda} = (\bar{\lambda}_k - \bar{\lambda}_l)_{j=1, j \sim (k,l)}^n$ for some $\bar{\lambda} = (\bar{\lambda}_i)_{i=1}^m$ and $v_i = A_i^T \mu_i$ for some $\mu_i \in \mathfrak{R}$, where $k \sim (i, j)$ means that column k has a 1 in row i and a -1 in row j or, equivalently, arc k is directed from node i to node j . Choose any spanning tree for the digraph and choose any node \bar{i} . Let $\lambda_i = \bar{\lambda}_i + (\mu_{\bar{i}} - \bar{\lambda}_{\bar{i}})$ and $u_i = A_i^T \lambda_i$ for all nodes i in the network. Since each node i can be reached from \bar{i} via a simple path P_i in the spanning tree, we have

$$|\lambda_i - \mu_i| = \left| - \sum_{(k,l) \in P_i^+} (\lambda_k - \mu_k - \lambda_l + \mu_l) + \sum_{(k,l) \in P_i^-} (\lambda_k - \mu_k - \lambda_l + \mu_l) \right|$$

The matrix A is symmetric and $A^2 = A$. The matrix A is idempotent. The matrix A is a projection matrix. The matrix A is a projection matrix. The matrix A is a projection matrix.

$$A^2 = A$$

We assume that A is a projection matrix. We assume that A is a projection matrix. We assume that A is a projection matrix. We assume that A is a projection matrix. We assume that A is a projection matrix.

$$A^2 = A$$

Let us show that for any $x \in \mathbb{R}^n$, we have $A(Ax) = Ax$.

$$A(Ax) = A^2x = Ax$$

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then $Ax = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$. Then $A(Ax) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = Ax$. This shows that $A(Ax) = Ax$.

$$A(Ax) = Ax$$

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Then $Ax = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$. Then $A(Ax) = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = Ax$.

$$A(Ax) = Ax$$

$$A(Ax) = Ax$$

The matrix A is a projection matrix. The matrix A is a projection matrix. The matrix A is a projection matrix. The matrix A is a projection matrix. The matrix A is a projection matrix.

$$A^2 = A$$

$$\begin{aligned}
&= \left| - \sum_{(k,l) \in P_i^+} (\bar{\lambda}_k - \bar{\lambda}_l - \mu_k + \mu_l) + \sum_{(k,l) \in P_i^-} (\bar{\lambda}_k - \bar{\lambda}_l - \mu_k + \mu_l) \right| \\
&\leq \sum_{(k,l) \in P_i} |\bar{\lambda}_k - \bar{\lambda}_l - \mu_k + \mu_l| \\
&\leq \sqrt{h_i} \left(\sum_{(k,l) \in P_i} |\bar{\lambda}_k - \bar{\lambda}_l - \mu_k + \mu_l|^2 \right)^{\frac{1}{2}} \\
&\leq \sqrt{h_i} \left(\sum_{\substack{j=1 \\ j \sim (k,l)}}^n |\bar{\lambda}_k - \bar{\lambda}_l - \mu_k + \mu_l|^2 \right)^{\frac{1}{2}} \\
&= \sqrt{h_i} \left\| \bar{u} - \sum_{p=1}^m v_p \right\|,
\end{aligned}$$

where P_i^+ and P_i^- denote the set of forward arcs and backward arcs in P_i and h_i denotes the number of arcs in P_i . Thus,

$$\sum_{i=1}^m \|u_i - v_i\|^2 = \sum_{i=1}^m \|A_i^T(\lambda_i - \mu_i)\|^2 = \sum_{i=1}^m \|A_i^T\|^2 |\lambda_i - \mu_i|^2 \leq \sum_{i=1}^m d_i h_i \left\| \bar{u} - \sum_{p=1}^m v_p \right\|^2,$$

where d_i is the number of arcs incident to node i . This shows that (5) holds with $C_1 = \sqrt{\sum_{i=1}^m d_i h_i}$. Notice that $\sum_{i=1}^m d_i = 2n$ and h_i is at most the diameter of the spanning tree. Since the choice of the spanning tree and the node \bar{i} are arbitrary, we can choose them to minimize C_1 . Also, $A_i A_j^T = 0$ if and only if nodes i and j are not joined by an arc, so $\hat{c} = \max\{d_1, \dots, d_m\}$ and the coloring of K_1, \dots, K_m is equivalent to graph coloring on the digraph.

In the above case of a nonlinear network flow problem, if G is also separable in the sense that $G(x) = \sum_{j=1}^n G_j(x_j)$ for all $x = (x_j)_{j=1}^n$ and $G_j : \mathfrak{R} \mapsto \mathfrak{R}$, then $\pi_i(u_1, \dots, u_m)$ given by (9) depends on only those u_k for which node k is a neighbor of node i and the asynchronous method (10)-(12) reduces to the asynchronous network relaxation method studied in [5, §7.2.3] and [38]. It is known that iterates generated by this method converge for any stepsize $\gamma \in (0, 1)$, assuming G^* is convex differentiable and (41) has a solution (G need not be defined everywhere on \mathfrak{R}^n and $(G^*)'$ need not be strongly monotone or Lipschitz continuous). However, no rate of convergence result was known. By applying Theorem 1, we obtain that this method has a linear rate of convergence, assuming $(G^*)'$ is strongly monotone and Lipschitz continuous and the stepsize is sufficiently small.

7 Applications to Partial Differential Equations

In this section we consider the Sobolev space $V = H_0^1(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega\}$ with duality pairing $\langle u, v \rangle = \int_{\Omega} (\sum_{i=1}^d \partial_i u \partial_i v + uv) dx$ and norm $\|v\| = \|v\|_{H^1(\Omega)} = \langle v, v \rangle^{\frac{1}{2}}$, where Ω is an open, bounded, and connected subset of \mathfrak{R}^d with Lipschitz continuous boundary $\partial\Omega$, $H^1(\Omega) = \{v \in L^2(\Omega) : \partial_i v \in L^2(\Omega), i = 1, \dots, d\}$, and $\partial_i v$ is the locally Lebesgue integrable real function defined on Ω satisfying $\int_{\Omega} \partial_i v \phi dx = - \int_{\Omega} v \frac{\partial \phi}{\partial x_i} dx$ for all $\phi \in C_0^\infty(\Omega) = \{\phi \in C^\infty(\Omega) : \phi \text{ has compact support}\}$ [15, pp. 10-13]. We will consider two nonlinear elliptic partial differential equations formulated as the minimization problem (2) and, for each, we will consider the space decomposition (4) corresponding to, respectively, DD and MG methods, and we will develop corresponding estimates for C_1 in (5), for C_2 in (6) and for c in (7)-(8). Throughout, we denote $|x| = (\sum_{i=1}^d x_i^2)^{\frac{1}{2}}$ for any $x = (x_i)_{i=1}^d \in \mathfrak{R}^d$.

The first partial differential equation corresponds to the minimization problem (2) with

$$K = H_0^1(\Omega), \quad \langle F'(u), v \rangle = \int_{\Omega} \left(\sum_{i=1}^d a_i(x, u, \nabla u) \partial_i v + a_0(x, u, \nabla u) v - f v \right) dx, \quad (44)$$

$$\begin{aligned}
 & \sum_{k=0}^{\infty} (k+1) x^k = \sum_{k=0}^{\infty} (k+1) x^k \\
 & \sum_{k=0}^{\infty} (k+1) x^k = \sum_{k=0}^{\infty} (k+1) x^k \\
 & \sum_{k=0}^{\infty} (k+1) x^k = \sum_{k=0}^{\infty} (k+1) x^k \\
 & \sum_{k=0}^{\infty} (k+1) x^k = \sum_{k=0}^{\infty} (k+1) x^k
 \end{aligned}$$

where $\sum_{k=0}^{\infty} x^k$ is the sum of the geometric series and $\sum_{k=0}^{\infty} (k+1) x^k$ is the sum of the series $\sum_{k=0}^{\infty} (k+1) x^k$.

$$\sum_{k=0}^{\infty} (k+1) x^k = \sum_{k=0}^{\infty} (k+1) x^k$$

where $\sum_{k=0}^{\infty} x^k$ is the sum of the geometric series and $\sum_{k=0}^{\infty} (k+1) x^k$ is the sum of the series $\sum_{k=0}^{\infty} (k+1) x^k$. This shows that the sum of the series $\sum_{k=0}^{\infty} (k+1) x^k$ is $\frac{1}{(1-x)^2}$. In fact, it can be shown that the sum of the series $\sum_{k=0}^{\infty} (k+1) x^k$ is $\frac{1}{(1-x)^2}$ for $|x| < 1$. This is done by differentiating the geometric series $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ with respect to x . The result is $\sum_{k=0}^{\infty} k x^{k-1} = \frac{1}{(1-x)^2}$. Multiplying both sides by x gives $\sum_{k=0}^{\infty} k x^k = \frac{x}{(1-x)^2}$. Adding $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ to both sides yields $\sum_{k=0}^{\infty} (k+1) x^k = \frac{x}{(1-x)^2} + \frac{1}{1-x} = \frac{x + 1 - x}{(1-x)^2} = \frac{1}{(1-x)^2}$.

7 Applications to Partial Differential Equations

In this section we consider the problem of solving the partial differential equation $\Delta u = f$ in a domain Ω with boundary $\partial\Omega$. The function u is assumed to be harmonic in Ω and to satisfy the boundary conditions $u = g$ on $\partial\Omega$. The method of images is used to solve this problem. The idea is to replace the domain Ω by a larger domain $\tilde{\Omega}$ in which the boundary is a sphere. The function u is then extended to $\tilde{\Omega}$ by the method of images. The resulting function \tilde{u} is harmonic in $\tilde{\Omega}$ and satisfies the boundary conditions on $\partial\tilde{\Omega}$. The function u is then recovered from \tilde{u} by restriction to Ω .

$$\Delta u = f \quad \text{in } \Omega$$

where $f \in L^2(\Omega)$ and $\nabla u = (\partial_i u)_{i=1}^d$ is the gradient of u [16, p. 302]. It is assumed that each nonlinear coefficient $a_i(x, p)$ is a real-valued function of $x = (x_j)_{j=1}^d$ and $p = (p_k)_{k=0}^d$ and is sufficiently smooth in the sense that

$$a_i \in C^1(\Omega \times \mathbb{R}^{d+1}), \quad (45)$$

$$\max_{\substack{j=1,2,\dots,d \\ k=0,1,\dots,d}} \left\{ |a_i(x, p)|, \left| \frac{\partial a_i}{\partial x_j}(x, p) \right|, \left| \frac{\partial a_i}{\partial p_k}(x, p) \right| \right\} \leq L, \quad (46)$$

for all $(x, p) \in \Omega \times \mathbb{R}^{d+1}$ and $i = 0, 1, \dots, d$, with L a constant. In addition, the matrix $\left[\frac{\partial a_i}{\partial p_k}(x, p) \right]_{i,k=0}^d$ is assumed to be uniformly positive definite, i.e.,

$$\sum_{i=0}^d \sum_{k=0}^d \frac{\partial a_i}{\partial p_k}(x, p) \xi_i \xi_k \geq \sigma \sum_{i=0}^d \xi_i^2, \quad \forall \xi_i \in \mathbb{R}, \quad i = 0, 1, \dots, d, \quad (47)$$

for all $(x, p) \in \Omega \times \mathbb{R}^{d+1}$, with $\sigma > 0$ a constant. Under these assumptions, the problem (2), which has the equation formulation

$$(F'(u), v) = 0, \quad \forall v \in H_0^1(\Omega), \quad (48)$$

is well posed and has a unique solution $u \in H_0^1(\Omega)$ (see [16, p. 302] and [21]). Moreover, straightforward calculation shows that

$$\begin{aligned} \langle F'(u) - F'(v), u - v \rangle &\geq \sigma \|u - v\|^2, \\ \langle F'(u) - F'(v), w \rangle &\leq L(d+1) \|u - v\| \|w\|, \end{aligned} \quad (49)$$

for all $u, v, w \in H^1(\Omega)$, so F' is strongly monotone and Lipschitz continuous.

The second partial differential equation corresponds to the minimization problem (2) with

$$K = H_0^1(\Omega), \quad F(v) = \int_{\Omega} \left(\frac{1}{2} |\nabla v|^2 + \frac{1}{4} v^4 - f v \right) dx, \quad (50)$$

where $f \in L^2(\Omega)$ and $d \in \{2, 3\}$. The corresponding equation is the simplified Ginzburg-Landau equation for superconductivity:

$$\begin{aligned} -\Delta u + u^3 &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \quad (51)$$

where u is the wave function, which is valid in the absence of internal magnetic field [36], and $\Delta u = \sum_{i=1}^d \partial_i(\partial_i u)$ denotes the Laplacian of u . Notice that F' has the form (44), with $a_0(x, p) = p_0^3$ and $a_i(x, p) = p_i$, $i = 1, \dots, d$, which does not satisfy (47). Nevertheless, straightforward calculation shows

$$\langle F'(u) - F'(v), u - v \rangle = \int_{\Omega} |\nabla u - \nabla v|^2 + (u^3 - v^3)(u - v) dx \geq \int_{\Omega} |\nabla u - \nabla v|^2 dx = \|u - v\|_{1,\Omega}^2,$$

for all $u, v \in H^1(\Omega)$. Since the semi-norm $|\cdot|_{1,\Omega}$ is equivalent to the norm $\|\cdot\|$ on $H_0^1(\Omega)$ [15, p. 12], this shows F' is strongly monotone on $H_0^1(\Omega)$.

In §7.1 and §7.2 below, we will study asynchronous DD and MG methods for solving the above two equations (48) and (51). We will analyze the convergence rate of the methods by estimating the constants C_1 , C_2 and c for the corresponding space decomposition of the finite element approximation subspace and then applying Theorem 1. In particular, we will show that the above two equations can be solved in parallel with a convergence factor that is independent of the finite element mesh size h , i.e., the number of iterations to reach a desired solution accuracy is independent of h .

where $\lambda = \lambda_1, \dots, \lambda_n$ and $\mu = \mu_1, \dots, \mu_n$ are the eigenvalues of A and B respectively. It is assumed that $\lambda_i \neq \mu_j$ for all i, j . The matrix M is defined by

$$M_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (1)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that M is invertible. The matrix N is defined by

$$N_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (2)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that N is invertible. The matrix P is defined by

$$P_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (3)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that P is invertible. The matrix Q is defined by

$$Q_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (4)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that Q is invertible. The matrix R is defined by

$$R_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (5)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that R is invertible. The matrix S is defined by

$$S_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (6)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that S is invertible. The matrix T is defined by

$$T_{ij} = \frac{1}{\lambda_i - \mu_j} \left(\frac{1}{\lambda_i} \delta_{ij} - \frac{1}{\lambda_i - \mu_j} \right) \quad (7)$$

for all $i, j \in \{1, \dots, n\}$. It is assumed that T is invertible. The matrix U is defined by

7.1 Domain decomposition methods

7.1.1 Decomposition of the domain Ω

In DD methods, the domain Ω is decomposed into the disjoint union of subdomains Ω_i , $i = 1, \dots, m$, 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 $i \neq j$. This is illustrated in Figure 1 where a rectangular-shaped domain in \mathbb{R}^2 is decomposed into the disjoint union of $m = 25$ rectangular-shaped subdomains and their boundary. The subdomains, which are assumed to form a regular quasi-uniform division (see p. 124 and Eq. (3.2.28) of [15] 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 (i.e., maximum diameter) 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 , $i = 1, \dots, m$, 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 r th-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 $i = 1, \dots, m$, let K_i denote the continuous, piecewise r th-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 $i = 0, 1, \dots, m$, and it can be shown that

$$K = \sum_{i=0}^m K_i,$$

so 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 $d = 3$. For the example of Figure 1, $d = 2$ and $n_c = 4$. Then the total number of colors needed for (7) and (8) to hold is $c = n_c + 1$.

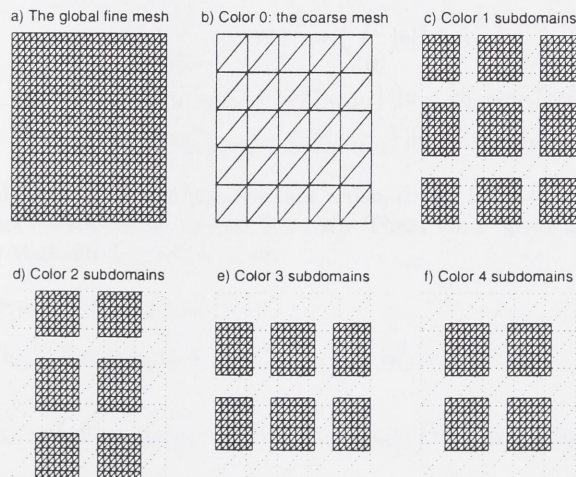


Figure 1: Decomposition of a rectangular-shaped domain in \mathbb{R}^2 .

7.1.2 Estimating C_1 for equations (48) and (51)

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 $i = 0, 1, \dots, m$ and satisfy $v = \sum_{i=0}^m v_i$ [29, pp. 163-165], [39, p. 607]. Moreover, by further choosing θ_i so that $|\nabla\theta_i|$ has a certain boundedness property, it was recently shown in [35, Lem. 4.1] that, for any $s \geq 1$,

$$\left(\sum_{i=0}^m \|v_i\|^s \right)^{\frac{1}{s}} \leq C c^{\frac{1}{s}} \left(1 + \left(\frac{H}{\delta} \right)^{\frac{1}{2}} \right) \|v\|,$$

where C is a constant independent of the mesh parameters and m . Taking $s = 2$ and using the subspace nature of K_i , we obtain that, for any $v_i \in K_i$, $i = 1, \dots, m$, there exists $\bar{u}_i \in K_i$ satisfying (5) (with summation index from 0 to m), where

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

Also see [12, Thm. 16] and a work of Dryja and Widlund cited therein for related results. By choosing the overlapping size δ proportional to the coarse-mesh size H , the constant C_1 will be independent of the mesh parameters and the number of subdomains m .

7.1.3 Estimating C_2 for equations (48) and (51)

Consider F given by (50), associated with the equation (51). By the mean value theorem, for any $u \in \mathfrak{R}$, $v \in \mathfrak{R}$, we have $|u^3 - v^3| = 3|\theta u + (1-\theta)v|^2|u-v| \leq 3(|u|+|v|)^2|u-v| \leq 6(|u|^2+|v|^2)|u-v|$ for some $\theta \in [0, 1]$. Thus, using the continuous embedding of $H^1(\Omega)$ in $L^p(\Omega)$ for $p < 2d/(d-2)$ and $d = 2, 3$ (see [15, p. 114], [19, p. 21]), we have for any $u, v \in H^1(\Omega)$ and any subdomain Ω' of Ω that $u, v \in L^4(\Omega)$ and

$$\begin{aligned} \left| \int_{\Omega'} (u^3 - v^3)w \, dx \right| &\leq 6 \int_{\Omega'} |u|^2|u-v||w| + |v|^2|u-v||w| \, dx \\ &\leq 6 \left(\left(\int_{\Omega'} |u|^4 \, dx \right)^{\frac{1}{2}} + \left(\int_{\Omega'} |v|^4 \, dx \right)^{\frac{1}{2}} \right) \left(\int_{\Omega'} |u-v|^2|w|^2 \, dx \right)^{\frac{1}{2}} \\ &\leq 6 \left(\|u\|_{L^4(\Omega')}^2 + \|v\|_{L^4(\Omega')}^2 \right) \|u-v\|_{L^4(\Omega')} \|w\|_{L^4(\Omega')} \\ &\leq C \left(\|u\|_{H^1(\Omega')}^2 + \|v\|_{H^1(\Omega')}^2 \right) \|u-v\|_{H^1(\Omega')} \|w\|_{H^1(\Omega')}, \end{aligned}$$

where C depends only on the embedding constant. Also, define $\Omega_0^\delta = \Omega$ for convenience, so that every $v \in K_i$ vanishes outside of Ω_i^δ ($i = 0, 1, \dots, m$). Then, for F given by (50), we have from the above inequality that, for $i, j = 0, 1, \dots, m$,

$$\begin{aligned} a_{ij} &= \langle F'(w_{ij} + u_{ij}) - F'(w_{ij}), v_i \rangle \\ &= \int_{\Omega_i^\delta \cap \Omega_j^\delta} (\nabla u_{ij})^T \nabla v_i + u_{ij} v_i + ((w_{ij} + u_{ij})^3 - w_{ij}^3) v_i \, dx \\ &\leq \left(1 + C \|w_{ij} + u_{ij}\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)}^2 + C \|w_{ij}\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)}^2 \right) \|u_{ij}\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)} \|v_i\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)}, \end{aligned} \quad (52)$$

for any $w_{ij} \in K$, $u_{ij} \in K_j$, $v_i \in K_i$, with $a_{ij} = 0$ whenever $\Omega_i^\delta \cap \Omega_j^\delta = \emptyset$. Assume there exists a constant $\alpha > 0$ such that $\|w_{ij} + u_{ij}\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)}^2 + \|w_{ij}\|_{H^1(\Omega_i^\delta \cap \Omega_j^\delta)}^2 \leq \alpha$ for $i, j = 0, 1, \dots, m$. Also, for $i, j = 1, \dots, m$, let $\epsilon_{ij} = 0$ if $\Omega_i^\delta \cap \Omega_j^\delta = \emptyset$ and otherwise let $\epsilon_{ij} = 1$. Let \hat{c} be the smallest integer such that every subdomain intersects at most \hat{c} other subdomains. It is not difficult to

7.1.2. Estimating β for equation (7.1) and (7.2)

Let $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$ be a $k \times 1$ vector of parameters. Let $X = (X_1, X_2, \dots, X_k)$ be a $n \times k$ matrix of explanatory variables. Let $Y = (Y_1, Y_2, \dots, Y_n)'$ be a $n \times 1$ vector of dependent variables. Then the linear regression model can be written as $Y = X\beta + \epsilon$, where $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)'$ is a $n \times 1$ vector of error terms. The least squares estimator of β is given by $\hat{\beta} = (X'X)^{-1}X'Y$.

$$\hat{\beta} = (X'X)^{-1}X'Y$$

where X' is a $k \times n$ matrix, $X'X$ is a $k \times k$ matrix, and $X'Y$ is a $k \times 1$ vector. The matrix $X'X$ is symmetric and positive definite. The vector $X'Y$ is the cross-product of the explanatory variables and the dependent variable.

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Let $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ be the elements of the vector $\hat{\beta}$. Then $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ are the least squares estimates of the parameters $\beta_1, \beta_2, \dots, \beta_k$. The variance-covariance matrix of $\hat{\beta}$ is given by $\text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$, where σ^2 is the variance of the error terms.

7.1.3. Estimating β for equation (7.3) and (7.4)

Let $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$ be a $k \times 1$ vector of parameters. Let $X = (X_1, X_2, \dots, X_k)$ be a $n \times k$ matrix of explanatory variables. Let $Y = (Y_1, Y_2, \dots, Y_n)'$ be a $n \times 1$ vector of dependent variables. Then the linear regression model can be written as $Y = X\beta + \epsilon$, where $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)'$ is a $n \times 1$ vector of error terms. The least squares estimator of β is given by $\hat{\beta} = (X'X)^{-1}X'Y$.

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Let $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ be the elements of the vector $\hat{\beta}$. Then $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ are the least squares estimates of the parameters $\beta_1, \beta_2, \dots, \beta_k$. The variance-covariance matrix of $\hat{\beta}$ is given by $\text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$, where σ^2 is the variance of the error terms.

$$\text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$$

Let $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ be the elements of the vector $\hat{\beta}$. Then $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ are the least squares estimates of the parameters $\beta_1, \beta_2, \dots, \beta_k$. The variance-covariance matrix of $\hat{\beta}$ is given by $\text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2$, where σ^2 is the variance of the error terms.

show that the symmetric matrix $\mathcal{E} = [\epsilon_{ij}]_{i,j=1}^m$ has the following estimate of its spectral radius (see [35, Corollary 5.1] for a proof):

$$\rho(\mathcal{E}) \leq \max_{i=1,\dots,m} \sum_{j=1}^m \epsilon_{ij} \leq \hat{c}.$$

This together with the estimate (52) yields

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^m a_{ij} &\leq (1 + C\alpha) \sum_{i=1}^m \sum_{j=1}^m \epsilon_{ij} \|u_{ij}\| \|v_i\| \\ &\leq (1 + C\alpha) \sum_{i=1}^m \sum_{j=1}^m \epsilon_{ij} \max_{i=1,\dots,m} \|u_{ij}\| \|v_i\| \\ &= (1 + C\alpha) \hat{c} \left(\sum_{j=1}^m \max_{i=1,\dots,m} \|u_{ij}\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Next, by using the fact Ω_j^δ , $j \in I(k)$, are disjoint subsets of Ω for $k = 1, \dots, c$, the estimate (52) yields

$$\begin{aligned} \sum_{j=1}^m a_{0j} &\leq (1 + C\alpha) \sum_{j=1}^m \|u_{0j}\| \|v_0\|_{H^1(\Omega_j^\delta)} \leq (1 + C\alpha) \left(\sum_{j=1}^m \|u_{0j}\|^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^m \|v_0\|_{H^1(\Omega_j^\delta)}^2 \right)^{\frac{1}{2}} \\ &\leq (1 + C\alpha) \sqrt{c} \left(\sum_{j=1}^m \|u_{0j}\|^2 \right)^{\frac{1}{2}} \|v_0\|, \quad \forall u_{0j} \in K_j, \forall v_0 \in K_0. \end{aligned}$$

Similar to the above argument, the estimate (52) gives

$$\begin{aligned} \sum_{i=1}^m a_{i0} &\leq (1 + C\alpha) \sum_{i=1}^m \|u_{i0}\|_{H^1(\Omega_i^\delta)} \|v_i\| \\ &\leq (1 + C\alpha) \left(\sum_{i=1}^m \|u_{i0}\|_{H^1(\Omega_i^\delta)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}}, \quad \forall u_{i0} \in K_0, \forall v_i \in K_i. \end{aligned}$$

We combine these estimates to obtain

$$\begin{aligned} \sum_{i=0}^m \sum_{j=0}^m a_{ij} &= a_{00} + \sum_{j=1}^m a_{0j} + \sum_{i=1}^m \sum_{j=1}^m a_{ij} + \sum_{i=1}^m a_{i0} \\ &\leq (1 + C\alpha) \|u_{00}\| \|v_0\| + (1 + C\alpha) \sqrt{c} \left(\sum_{j=1}^m \|u_{0j}\|^2 \right)^{\frac{1}{2}} \|v_0\| \\ &\quad + (1 + C\alpha) \hat{c} \left(\sum_{j=1}^m \max_{i=1,\dots,m} \|u_{ij}\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}} \\ &\quad + (1 + C\alpha) \left(\sum_{i=1}^m \|u_{i0}\|_{H^1(\Omega_i^\delta)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}} \\ &\leq \tilde{C}_2 \left(\sum_{j=0}^m \max_{i=0,1,\dots,m} \|u_{ij}\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=0}^m \|v_i\|^2 \right)^{\frac{1}{2}} \\ &\quad + (1 + C\alpha) \left(\sum_{i=1}^m \|u_{i0}\|_{H^1(\Omega_i^\delta)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|v_i\|^2 \right)^{\frac{1}{2}}, \end{aligned} \tag{53}$$

show that the sequence $\sum_{k=1}^n \frac{1}{k^2}$ is bounded and hence convergent (use the Cauchy criterion).

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

This together with the estimate (1) yields

$$\sum_{k=1}^n \frac{1}{k^2} \leq 1 + \sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$

Just by using the fact $\sum_{k=1}^n \frac{1}{k^2} < 2$ we have $\sum_{k=1}^n \frac{1}{k^2} < 2$ for all $n \in \mathbb{N}$.

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

Since in the above argument the estimate (1) gives

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$

We conclude that estimate (1) yields

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\sum_{k=1}^n \frac{1}{k^2} < 2$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$

$$\left(\sum_{k=1}^n \frac{1}{k^2} \right)^2 < 4$$



with \tilde{C}_2 a constant depending on $C\alpha, c, \hat{c}$ only. Compared with (6) (with $i, j = 0, 1, \dots, m$), we see that (53) has an extra term on the right-hand side. In the appendix, we will show that this extra term does not affect the convergence rate result of §4. In particular, we will show that Lemmas 1 and 2 hold with $C_2 = \tilde{C}_2 + (1 + C\alpha)\sqrt{c}$, so that Theorem 1 is still valid.

For F specified by (44), associated with the equation (48), it can be similarly proved using (49) that (53) holds, possibly with different constants C and α .

Upon applying the asynchronous method (10)–(12) with the above choice of space decomposition and under the assumptions (13)–(14), we obtain a parallel DD method for (48) and (51) whose convergence factor, according to Theorem 1 and the above estimates of C_1 and C_2 and assuming the overlapping size δ is proportional to the coarse mesh size H , is independent of the mesh parameters and the number of the subdomains.

7.2 Multigrid methods

7.2.1 Construction of the multigrid subspaces

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 $\mathcal{T}_k, k = 1, \dots, J$, 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 [15, Eq. (3.2.28)] 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} .

For example, in the two-dimensional case of $d = 2$, if we construct \mathcal{T}_k by connecting the midpoints of the edges of the triangles of \mathcal{T}_{k-1} , with \mathcal{T}_1 being the given coarsest initial triangulation, the resulting sequence of triangulation is quasi-uniform and $r = 1/\sqrt{2}$ (see Figure 2). Corresponding to each triangulation \mathcal{T}_k , we define the finite element subspace:

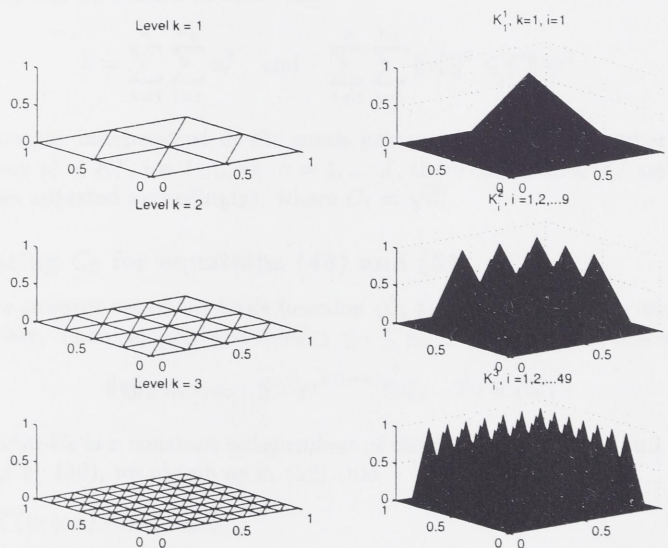


Figure 2: The multigrid mesh and basis functions.

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

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

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

with a constant depending on ϵ and δ . The dependence on δ is such that the
 as $\delta \rightarrow 0$ the error term on the right-hand side of the equation we get goes to zero
 since δ is not affected by the convergence rate of δ . In particular, we get that
 Lemma 1 and 2 hold with $C_1 = C_2 = (1 + C_0)/\delta$, so that Lemma 1 is still valid.
 For δ specified by (34), associated with the equation (31), it can be similarly verified that
 (35) and (36) hold jointly with different constants C and δ .
 Thus, applying the arguments worked (17)-(22) with the above choice of δ , Lemma
 position and under the assumptions (15)-(16), we obtain a general (30) needed for (25) and
 (24) where convergence factor involving in Lemma 1 and the above constants of C_1 and C_2
 and assuming the convergence rate δ is proportional to the error with size δ is independent
 of the mesh parameters and the number of the subdomains.

7.2. Multigrid methods

7.2.1. Construction of the multigrid algorithm

In this section it is desired to describe the multigrid algorithm \mathcal{M} for a constant coefficient
 linear elliptic problem on the domain $\Omega \subset \mathbb{R}^2$ with $\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_N$ as a grid
 composed of regular quadrilateral subdomains, and \mathcal{M} is a reduction of standard \mathcal{M} (see [15])
 of the (i.e. constant coefficient) \mathcal{M} with the $\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_N$ and the subdomains
 constants are independent of ϵ , i.e. \mathcal{M} is ϵ -independent and with each subdomain in Ω , using the same
 constants are independent of ϵ . The first reason that there is a constant $\nu < 1$ independent of ϵ and
 that \mathcal{M} is proportional to ν .
 For example, in the two-dimensional case of $\epsilon \ll 1$, if we consider \mathcal{M} for computing the
 eigenvalues of the matrix of \mathcal{M} , ν with \mathcal{M} being the given constant initial error
 reduction, the resulting eigenvalues of \mathcal{M} are given by ν and ν^2 (see Figure 2).
 Corresponding to each eigenvalue ν , we define the corresponding subdomain.



Figure 2. The spectral gap and error reduction.

$$\lambda_1 - \lambda_2 = (1 - \nu) \epsilon_k \quad (37)$$

where λ_1 (λ_2) denotes the eigen of the reduced matrix function of \mathcal{M} and ϵ_k denotes the error
 reduction with \mathcal{M} a total error (denoted by ϵ_k) that satisfies $\epsilon_k = \nu^k \epsilon$.
 \mathcal{M} is ϵ -independent.

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 $c = n_c J$.

7.2.2 Estimating C_1 for equations (48) and (51)

Let Q_k be the projection in the L^2 -norm onto the subspace \mathcal{M}_k , which is well defined on $H_0^1(\Omega) \subset L^2(\Omega)$. For any $v \in K$, let $v^k = (Q_k - Q_{k-1})v$, $k = 1, \dots, J$. Then, by Prop. 8.6 in [39, p. 611], we have

$$\sum_{k=1}^J \|v^k\|^2 \leq C_0 \|v\|^2,$$

where C_0 is a constant independent of the mesh parameters and J . By further decomposing each v^k as

$$v^k = \sum_{i=1}^{n_k} v_i^k \quad \text{with} \quad v_i^k = v^k(x_i^k) \phi_i^k,$$

the above estimate can be refined to show that

$$v = \sum_{k=1}^J \sum_{i=1}^{n_k} v_i^k \quad \text{and} \quad \sum_{k=1}^J \sum_{i=1}^{n_k} \|v_i^k\|^2 \leq C \|v\|^2,$$

where C is a constant independent of the mesh parameters and the number of levels J [35, §5.1]. Thus, for any $v_i^k \in K_i^k$, $i = 1, \dots, n_k$, $k = 1, \dots, J$, there exists $\bar{u}_i^k \in K_i^k$ satisfying (5) (with summation indices adjusted accordingly), where $C_1 = \sqrt{C}$.

7.2.3 Estimating C_2 for equations (48) and (51)

Let Λ_i^k denote the support set of the basis function ϕ_i^k , for all i and k . Also, recall the constant $r < 1$ defined earlier. Then, for any $k < l$ and $1 \leq i \leq n_k$, $1 \leq j \leq n_l$, the following estimate

$$\|u\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)} \leq C_0 r^{d(l-k)} \|u\|, \quad \forall u \in K_i^k,$$

can be shown, where C_0 is a constant independent of the mesh parameters and J [35, Eq. (56)]. Then, for F given by (50), we obtain as in (52) that

$$\begin{aligned} & \langle F'(w+u) - F'(w), v \rangle \\ & \leq \left(1 + C \|w+u\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)}^2 + C \|w\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)}^2 \right) \|u\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)} \|v\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)} \\ & \leq \left(1 + C \|w+u\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)}^2 + C \|w\|_{H^1(\Lambda_i^k \cap \Lambda_j^l)}^2 \right) C_0 r^{d(l-k)} \|u\| \|v\|, \end{aligned} \quad (54)$$

$$\forall w \in K, u \in K_i^k, v \in K_j^l,$$

where C is the embedding constant. For any i, j, k, l , defining

$$\varepsilon_{i,j}^{k,l} = \begin{cases} C_0 \gamma^{d|l-k|}, & \text{if } \text{supp}(\phi_i^k) \cap \text{supp}(\phi_j^l) \neq \emptyset; \\ 0, & \text{otherwise.} \end{cases}$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the matrix A . For each λ_i we have

$$\lambda_i = \frac{1}{n} \sum_{j=1}^n a_{ij}$$

Thus, $\lambda_1, \dots, \lambda_n$ are the following eigenvalues:

$$\lambda = \frac{1}{n} \sum_{i=1}^n a_{ii}, \quad \lambda = \frac{1}{n} \sum_{i=1}^n a_{ii}$$

The next step is to show that the matrix A is a real symmetric matrix. The number of rows and columns is a constant independent of the size n . The number of rows and columns is a constant independent of the size n . Thus the total number of rows and columns is a constant independent of the size n .

3.3. Estimating C for equations (48) and (51)

Let C be the projection in the L^2 norm onto the subspace \mathcal{H} . Then it will follow from (48) and (51) that C is a constant independent of the size n .

$$\sum_{i=1}^n |a_{ii}| \leq C$$

where C is a constant independent of the size n . Thus, C is a constant independent of the size n .

$$\lambda = \frac{1}{n} \sum_{i=1}^n a_{ii} \leq C$$

the above estimate can be used to show that

$$\sum_{i=1}^n |a_{ii}| \leq C, \quad \sum_{i=1}^n |a_{ii}| \leq C$$

where C is a constant independent of the size n . Thus, C is a constant independent of the size n .

3.4. Estimating C for equations (52) and (53)

Let C be the projection in the L^2 norm onto the subspace \mathcal{H} . Then it will follow from (52) and (53) that C is a constant independent of the size n .

$$\sum_{i=1}^n |a_{ii}| \leq C, \quad \sum_{i=1}^n |a_{ii}| \leq C$$

can be shown that C is a constant independent of the size n . Thus, C is a constant independent of the size n .

$$\lambda = \frac{1}{n} \sum_{i=1}^n a_{ii} \leq C$$

$$\sum_{i=1}^n |a_{ii}| \leq C, \quad \sum_{i=1}^n |a_{ii}| \leq C$$

$$\sum_{i=1}^n |a_{ii}| \leq C, \quad \sum_{i=1}^n |a_{ii}| \leq C$$

where C is the bounding constant. For any $\lambda \in \mathbb{R}$, we have

$$\lambda = \frac{1}{n} \sum_{i=1}^n a_{ii} \leq C$$

Assuming there exists a constant $\alpha > 0$ such that $\|w_{i,j}^{k,l} + u_{i,j}^{k,l}\|^2 + \|w_{i,j}^{k,l}\|^2 \leq \alpha$ for all i, j, k, l , the estimate (54) then yields

$$\begin{aligned} & \sum_{k=1}^J \sum_{i=1}^{n_k} \sum_{l=1}^J \sum_{j=1}^{n_l} \langle F'(w_{i,j}^{k,l} + u_{i,j}^{k,l}) - F'(w_{i,j}^{k,l}), v_i^k \rangle \\ & \leq C_0(1 + C\alpha) \sum_{i,k} \sum_{j,l} \varepsilon_{i,j}^{k,l} \|u_{i,j}^{k,l}\| \|v_i^k\| \\ & \leq C_0(1 + C\alpha) \sum_{k=1}^J \sum_{i=1}^{n_k} \sum_{l=1}^J \sum_{j=1}^{n_l} \varepsilon_{i,k}^{k,l} \max_{i,k} \|u_{i,j}^{k,l}\| \cdot \|v_i^k\|, \quad \forall u_{i,j}^{k,l} \in K_j^l, \forall v_i^k \in K_i^k. \end{aligned}$$

With proper ordering of the indices, the matrix $\mathcal{E} = [\varepsilon_{i,j}^{k,l}]$ is symmetric and its spectral radius $\rho(\mathcal{E})$ has been shown to be less than a constant independent of the mesh parameters and the number of levels [29, pp. 182–184]. Therefore

$$\begin{aligned} & \sum_{k=1}^J \sum_{i=1}^{n_k} \sum_{l=1}^J \sum_{j=1}^{n_l} \langle F'(w_{i,j}^{k,l} + u_{i,j}^{k,l}) - F'(w_{i,j}^{k,l}), v_i^k \rangle \\ & \leq C_0(1 + C\alpha)\rho(\mathcal{E}) \left(\sum_{l=1}^J \sum_{j=1}^{n_l} \max_{i,k} \|u_{i,j}^{k,l}\|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^J \sum_{i=1}^{n_k} \|v_i^k\|^2 \right)^{\frac{1}{2}}, \end{aligned}$$

which shows that (6) holds, with the constant $C_2 = C_0(1 + C\alpha)\rho(\mathcal{E})$ independent of the mesh parameters and the number of levels for the MG approximation.

For F specified by (44), it can be similarly proved that (6) holds with C_2 some constant independent of the mesh parameters and the number of levels.

Upon applying the asynchronous method (10)–(12) with the above choice of space decomposition and under the assumptions (13)–(14), we obtain a parallel MG method for (48) and (51) whose convergence factor, according to the above estimates of C_1 and C_2 and Theorem 1, is independent of the mesh parameters. This method generalizes the BPX multigrid method proposed in [8], which was used as a preconditioner for linear elliptic problems. Here, the parallel MG method is used as a solver and is applicable not only to linear, but also to nonlinear elliptic problems. And it further allows for asynchronous updates.

8 Appendix

In this appendix, we show that (53) can be used in place of (6) to prove Lemmas 1 and 2 for the DD method of §7.1, with $C_2 = \tilde{C}_2 + (1 + C\alpha)\sqrt{c}$. Here, the indices i and j are understood to always range over $0, 1, \dots, m$, instead of $1, \dots, m$.

First, we note that condition (6) is used only to show (22), (30) and (34) in the proofs. For (22), if we use condition (53) instead of (6), then (22) would have \tilde{C}_2 in place of C_2 and would have the following extra term on its right-hand side:

$$E = (1 + C\alpha)\gamma \left(\sum_{i=1}^m \|u_0(\tau_0^i(t)) - u_0(t)\|_{H^1(\Omega_i^t)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(t)\|^2 \right)^{\frac{1}{2}}.$$

Correspondingly, (24) would have \tilde{C}_2 in place of C_2 and would have the above extra term on its right-hand side. Using (23) and the fact that Ω_i^t , $i \in I(k)$, are disjoint subsets of Ω for $k = 1, \dots, c$, we see that

$$\begin{aligned} E & \leq (1 + C\alpha)\gamma^2 \sqrt{B} \left(\sum_{i=1}^m \sum_{\tau=t-B+1}^{t-1} \|s_0(\tau)\|_{H^1(\Omega_i^t)}^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(t)\|^2 \right)^{\frac{1}{2}} \\ & \leq (1 + C\alpha)\gamma^2 \sqrt{Bc} \left(\sum_{\tau=t-B+1}^{t-1} \|s_0(\tau)\|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \|s_i(t)\|^2 \right)^{\frac{1}{2}}, \end{aligned}$$

Assume that \mathbf{w} is a unit vector in \mathbb{R}^n and $\mathbf{w}^T \mathbf{w} = 1$. Then $\mathbf{w}^T \mathbf{A} \mathbf{w}$ is a scalar and $\mathbf{w}^T \mathbf{A} \mathbf{w} = \mathbf{w}^T (\mathbf{A} \mathbf{w})$.

$$\mathbf{w}^T \mathbf{A} \mathbf{w} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j a_{ij} = \sum_{i=1}^n \sum_{j=1}^n w_j w_i a_{ji} = \sum_{j=1}^n \sum_{i=1}^n w_j w_i a_{ji} = \mathbf{w}^T \mathbf{A}^T \mathbf{w}$$

With proper ordering of the entries, the matrix \mathbf{A} is symmetric and its eigenvalues are real. It has been shown in the text that a real symmetric matrix has n real eigenvalues and n orthogonal eigenvectors.

$$\mathbf{w}^T \mathbf{A} \mathbf{w} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j a_{ij} = \sum_{i=1}^n \sum_{j=1}^n w_j w_i a_{ji} = \sum_{j=1}^n \sum_{i=1}^n w_j w_i a_{ji} = \mathbf{w}^T \mathbf{A}^T \mathbf{w}$$

which shows that \mathbf{A} is Hermitian with the property $\mathbf{A} = \mathbf{A}^T$. It is independent of the choice of parameters and the number of terms in the MD approximation.

For a general $n \times n$ matrix \mathbf{A} , it can be shown that $\mathbf{A} + \mathbf{A}^T$ is symmetric and its eigenvalues are real.

To compute the eigenvalues and eigenvectors of \mathbf{A} , we use the method of Lagrange multipliers and solve the equations $\mathbf{A} \mathbf{w} = \lambda \mathbf{w}$ and $\mathbf{w}^T \mathbf{w} = 1$. The method of Lagrange multipliers is used to solve the constrained optimization problem. The method involves the use of the method of Lagrange multipliers. The method involves the use of the method of Lagrange multipliers. The method involves the use of the method of Lagrange multipliers.

Appendix 8

In this appendix, we show that the method of Lagrange multipliers can be used to solve the constrained optimization problem. The method involves the use of the method of Lagrange multipliers. The method involves the use of the method of Lagrange multipliers. The method involves the use of the method of Lagrange multipliers.

$$\mathbf{w}^T \mathbf{A} \mathbf{w} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j a_{ij} = \sum_{i=1}^n \sum_{j=1}^n w_j w_i a_{ji} = \sum_{j=1}^n \sum_{i=1}^n w_j w_i a_{ji} = \mathbf{w}^T \mathbf{A}^T \mathbf{w}$$

Consequently, \mathbf{A} is Hermitian and its eigenvalues are real. It has been shown in the text that a real symmetric matrix has n real eigenvalues and n orthogonal eigenvectors.

$$\mathbf{w}^T \mathbf{A} \mathbf{w} = \sum_{i=1}^n \sum_{j=1}^n w_i w_j a_{ij} = \sum_{i=1}^n \sum_{j=1}^n w_j w_i a_{ji} = \sum_{j=1}^n \sum_{i=1}^n w_j w_i a_{ji} = \mathbf{w}^T \mathbf{A}^T \mathbf{w}$$

which implies that (24) holds with $C_2 = \tilde{C}_2 + (1 + C\alpha)\sqrt{c}$. The remainder of the proof of Lemma 1 then proceeds as before.

For (30) and (34), a similar argument can be applied to show that Lemma 2 holds with the above choice of C_2 .

References

- [1] Z.-Z. Bai, V. Migallón, J. Penadés, and D. B. Szyld. Block and asynchronous two-stage methods for mildly nonlinear systems. *Num. Math.*, to appear. See also <http://www.math.temple.edu/~szyld>.
- [2] R. Bank. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39:453–465, 1982.
- [3] R. Bank and D. J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39:453–465, 1982.
- [4] D. P. Bertsekas, D. Castañón, J. Eckstein, and S. A. Zenios. Parallel computing in network optimization. In M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser, editors, *Handbooks in Operations Research and Management Science 7: Network Models*, pages 331–399, Amsterdam, 1995. Elsevier Scientific.
- [5] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Englewood Cliffs, 1989.
- [6] D. P. Bertsekas and J. N. Tsitsiklis. Some aspects of parallel and distributed iterative algorithms—a survey. *Automatica*, 27:3–21, 1991.
- [7] J. Bramble and X. Zhang. *The Analysis of Multigrid Methods*. North-Holland, to appear.
- [8] J. H. Bramble, J. E. Pasciak, and J.-C. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55:1–22, 1990.
- [9] A. Brandt. Multilevel adaptive solutions to boundary value problems. *Math. Comp.*, 31:333–309, 1977.
- [10] R. Bru, V. Migallón, J. Penadés, and D. B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Trans. Numer. Anal.*, 3:24–38, 1995.
- [11] J. V. Burke and P. Tseng. A unified analysis of Hoffman’s bound via Fenchel duality. *SIAM J. Optim.*, 6:265–282, 1996.
- [12] T. F. Chan and T. P. Mathew. Domain decomposition algorithms. *Acta Numerica*, pages 61–143, 1994.
- [13] T. F. Chan and I. Sharapov. Subspace correction multilevel methods for elliptic eigenvalue problems. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *Proceedings of the 9th International Domain Decomposition Conference*. John Wiley and Sons, to appear.
- [14] D. Chazan and W. L. Miranker. Chaotic relaxation. *Lin. Algeb. Appl.*, 2:199–222, 1969.
- [15] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [16] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, 37:296–311, 1997.
- [17] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North-Holland, Amsterdam, 1976.
- [18] A. Frommer, H. Schwardt, and D. B. Szyld. Asynchronous weighted additive Schwarz methods. *Electronic Trans. Numer. Anal.*, 5:48–61, 1997.
- [19] R. Glowinski and P. Le Tallec. *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM, Philadelphia, 1989.
- [20] W. Hackbusch and A. Reusken. Analysis of a damped nonlinear multilevel method. *Numer. Math.*, 55:225–246, 1989.

which includes the (20) ... The ... of the ...
... (20) ... a ... agreement ...
... (20) ...

References

- [1] ... and ...
...
- [2] ... and ...
...
- [3] ... and ...
...
- [4] ... and ...
...
- [5] ... and ...
...
- [6] ... and ...
...
- [7] ... and ...
...
- [8] ... and ...
...
- [9] ... and ...
...
- [10] ... and ...
...
- [11] ... and ...
...
- [12] ... and ...
...
- [13] ... and ...
...
- [14] ... and ...
...
- [15] ... and ...
...
- [16] ... and ...
...
- [17] ... and ...
...
- [18] ... and ...
...
- [19] ... and ...
...
- [20] ... and ...
...

- [21] O. A. Ladyzhenskaya and N. N. Uraltseva. *Linear and Quasilinear Elliptic Equations*. Academic Press, New York, 1968.
- [22] 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.
- [23] 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.
- [24] S. F. McCormick. *Multilevel Adaptive Methods for Partial Differential Equations*, volume 6 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1989.
- [25] R. Rannacher. On the convergence of the Newton-Raphson method for strongly nonlinear FE-equations. In P. Wriggers and W. Wagner, editors, *Nonlinear computational mechanics*, pages 11–30, Berlin, 1991. Springer-Verlag.
- [26] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.
- [27] R. T. Rockafellar. *Network Flows and Monotropic Programming*. Wiley-Interscience, New York, 1984.
- [28] I. A. Sharapov. Multilevel subspace correction for large scale optimization problems. Technical Report CAM-97-31, UCLA, Computational and Applied Mathematics, 1997.
- [29] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain Decomposition: Parallel Multilevel Algorithms for Elliptic Partial Differential Equations*. Cambridge University Press, Cambridge, 1996.
- [30] 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, Technische Universität Graz*, 1992. See also <http://www.mi.uib.no/~tai>.
- [31] 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, New York, 1994. Marcel Dekker. See also <http://www.mi.uib.no/~tai>.
- [32] 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, Pennsylvania State University, 1993)*, pages 355–360, Providence, 1995. American Mathematical Society.
- [33] X.-C. Tai. Parallel function decomposition and space decomposition methods: Part II. Space decomposition. *Beijing Mathematics*, 1, part 2:135–152, 1995. See also <http://www.mi.uib.no/~tai>.
- [34] 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.
- [35] X.-C. Tai and J.-C. Xu. Global convergence of subspace correction methods for convex optimization problems. Technical Report No.-114, Department of Mathematics, University of Bergen, Norway, March, 1998. See also <http://www.mi.uib.no/~tai>.
- [36] M. Tinkham. *Introduction to Superconductivity*. McGraw Hill, New York, 1975.
- [37] P. Tseng. On the rate of convergence of a partially asynchronous gradient projection algorithm. *SIAM J. Optim.*, 1:603–619, 1991.
- [38] P. Tseng, D. P. Bertsekas, and J. N. Tsitsiklis. Partially asynchronous, parallel algorithms for network flow and other problems. *SIAM J. Control Optim.*, 28:678–710, 1990.
- [39] J.-C. Xu. Iteration methods by space decomposition and subspace correction. *SIAM Rev.*, 34:581–613, 1992.
- [40] J.-C. Xu. A novel two-grid method for semilinear elliptic equations. *SIAM J. Sci. Comp.*, 15:231–237, 1994.
- [41] J.-C. Xu. Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.*, 33:1759–1777, 1996.

[1] O. A. Lachmann and A. K. Lubars, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1952.

[2] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[3] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[4] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[5] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[6] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[7] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[8] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[9] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[10] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[11] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[12] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[13] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[14] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[15] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

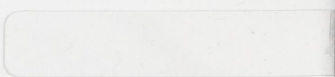
[16] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

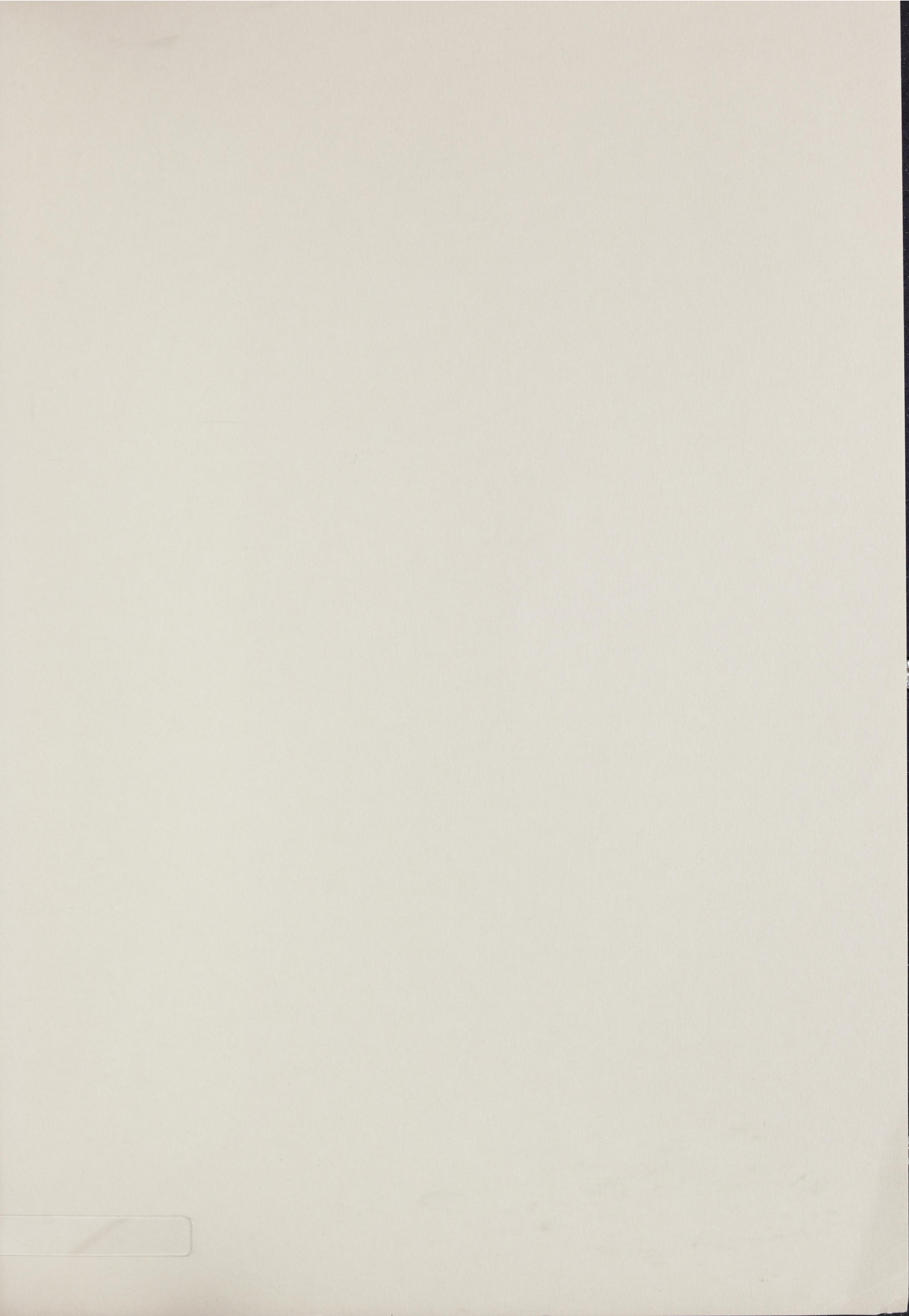
[17] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[18] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[19] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.

[20] E. D. Lashin, *Long-time behavior of solutions of elliptic equations*, *Journal of Mathematical Analysis and Applications*, **24**(1), 1978, pp. 1-15.







Depotbiblioteket



78sd 20 231

