

DISCRETE ANALYSIS OF DOMAIN DECOMPOSITION APPROACHES FOR MESH GENERATION VIA THE EQUIDISTRIBUTION PRINCIPLE

RONALD D. HAYNES* AND FELIX KWOK†

Abstract. Moving mesh methods based on the equidistribution principle are powerful techniques for the space-time adaptive solution of evolution problems. Solving the resulting coupled system of equations, namely the original PDE and the mesh PDE, however, is challenging in parallel. Recently several Schwarz domain decomposition algorithms were proposed for this task and analyzed at the continuous level. However, after discretization, the resulting problems may not even be well posed, so the discrete algorithms requires a different analysis, which is the subject of this paper. We prove that when the number of grid points is large enough, the classical parallel and alternating Schwarz methods converge to the unique monodomain solution. Thus, such methods can be used in place of Newton’s method, which can suffer from convergence difficulties for challenging problems. The analysis for the nonlinear domain decomposition algorithms is based on M -function theory and is valid for an arbitrary number of subdomains. An asymptotic convergence rate is provided and numerical experiments illustrate the results.

Key words. domain decomposition, Schwarz methods, moving meshes, equidistribution, discretization, M -functions

AMS subject classifications. 65M55, 65N22, 65Y05, 65M50, 65N50

1. Introduction. The numerical simulation of partial differential equations (PDEs) can often benefit from the use of non-uniform meshes chosen to adapt to the local solution structures, especially in cases where the solutions vary over disparate space and/or time scales. In this article, we consider a class of r -refinement or moving mesh methods. This PDE-based mesh generation approach adapts an initial (typically uniform) grid by *relocating* a fixed number of mesh nodes while preserving the mesh topology. The mesh is determined by solving a so-called *moving mesh PDE* (*MMPDE*) which is coupled to the physical PDE of interest.

In its simplest form, the r -refinement technique is guided in large part by the equidistribution principle (EP) of de Boor [15] and Burchard [7]. In one dimension, the EP seeks a mesh so that the error contribution on each interval is the same. This naturally places mesh points closer together in regions of the physical space Ω_p where the error in the solution $u(t, x)$ is large. Mathematically, given some positive measure of the error, $M(t, x, u)$, for $x \in \Omega_p$ and $t \in [0, T]$, the EP requires that the mesh points $x_i, i = 0, \dots, N$, satisfy

$$\int_{x_{i-1}}^{x_i} M(t, \tilde{x}, u) d\tilde{x} \equiv \frac{1}{N} \int_0^1 M(t, \tilde{x}, u) d\tilde{x}.$$

The function M is referred to as the monitor or mesh density function.

In the continuous setting, we seek a time dependent mesh transformation between an underlying computational coordinate $\xi \in \Omega_c$ and the physical coordinate, $x \in \Omega_p$, so that

$$\int_0^{x(\xi_i, t)} M(t, \tilde{x}, u) d\tilde{x} = \frac{i}{N} \theta(t) \equiv \xi_i \theta(t), \quad (\text{EP})$$

*Department of Mathematics & Statistics, Memorial University of Newfoundland, St. John’s, Newfoundland, Canada, A1C 5S7

†Section de mathématiques, Université de Genève, 2-4 Rue du Lièvre, CP 64, 1211 Geneva 4, Switzerland

where $\theta(t) \equiv \int_0^1 M(t, \tilde{x}, u) d\tilde{x}$ is the total error in the solution. It follows directly, by differentiating the continuous form of (EP), that the required mesh transformation satisfies, for all t , the nonlinear differential equation

$$\frac{\partial}{\partial \xi} \left\{ M(t, x(\xi, t), u) \frac{\partial}{\partial \xi} x(\xi, t) \right\} = 0. \quad (1.1)$$

For time dependent problems the EP is often relaxed [30], and the mesh may be found by solving

$$\frac{\partial x}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial \xi} \left(M(t, x(\xi, t), u) \frac{\partial x}{\partial \xi} \right). \quad (\text{MMPDE5})$$

The relaxation parameter τ is chosen in practice (see [29]) so that the mesh evolves at a rate commensurate with that of the physical solution $u(t, x)$. A parabolic mesh equation provides a degree of temporal smoothing for the mesh generation problem. Coupling (MMPDE5) with a time dependent physical PDE and discretizing in space gives a initial value problem to solve. In contrast, coupling (1.1) with a physical PDE and discretizing in space gives an often more difficult DAE problem. The well-posedness of the steady mesh generation problem (1.1) and the semi-discretized (in time) form of (MMPDE5) may be found in [22]. A thorough treatise of r -refinement methods may be found in the recent book [32].

The mesh generators (1.1) and (MMPDE5) depend on the physical solution u which itself is the unknown solution of a physical PDE. In practice, this coupled system of DEs may be solved using the *MP* iterative procedure [32]. An initial guess for the mesh is chosen and the physical PDE is solved to give an approximate solution. This mesh and physical solution can be used to evaluate M and then (1.1) or (MMPDE5) can be solved for the new mesh, and the process may be repeated. This is in contrast to the approach, taken in MOVCOL [31] for example, which solves for the mesh and physical solution as one large simultaneous system.

Recently we have been interested in the parallel solution of PDEs on (moving) equidistributing grids. Our solution strategy introduces spatial parallelism in both the computation of the underlying meshes and the physical PDE by a domain decomposition (DD) approach. DD solves a PDE by divide and conquer: the spatial domain Ω is partitioned into overlapping or non-overlapping subdomains, thus reformulating the original PDE into an equivalent coupled system of PDEs which is solved by iteration. The coupling is provided by transmission conditions which are designed to appropriately match solutions on neighboring subdomains. The matching is imposed in a way which accelerates the convergence of the DD iteration and to maintain as much coarse grain granularity in the computation as possible. This allows implementation on a distributed memory architecture using MPI (for example). A stochastic DD approach for mesh generation in 2D which does not require iteration has recently been considered in [4].

Another benefit of the DD approach is its robustness relative to other nonlinear solution schemes. When the monitor function M is “difficult” (e.g. highly oscillatory), Newton’s method often fails to converge when applied to the discretized versions of (EP) and (MMPDE5). Then it is natural to divide the interval $[0, 1]$ into several subdomains to obtain smaller, “easier” problems, and use a DD iteration to obtain the global solution. We will illustrate this approach in Section 6.3.

In [22], various parallel Schwarz algorithms were presented to solve (1.1) and the semi-discretized (in time) form of (MMPDE5) using domain decomposition in the

computational co-ordinate ξ . These algorithms were analyzed at the continuous level, were shown to converge, and contraction rates obtained. Alternating algorithms, which give faster convergence at the cost of losing immediate parallelization, were presented in [28]. These analyses assume M is a function of the required mesh only, as is the case in the MP procedure mentioned above.

In practice these Schwarz methods have to be discretized for implementation. For instance, (1.1) can be discretized as

$$\begin{aligned} \tilde{M}(x_i, x_{i+1})(x_{i+1} - x_i) - \tilde{M}(x_{i-1}, x_i)(x_i - x_{i-1}) &= 0, \quad i = 1, \dots, N-1, \\ x_0 = x_L, \quad x_N = x_R, \end{aligned} \quad (1.2)$$

where $x_L = x(0)$ and $x_R = x(1)$ are the boundary conditions, and the quantity $\tilde{M}(u, v)$ is chosen so that (1.2) is a consistent discretization of (1.1). For example, $\tilde{M}(u, v) = M((u+v)/2)$ and $\tilde{M}(u, v) = (M(u) + M(v))/2$ gives the midpoint and trapezoidal discretizations respectively.

Likewise, we discretize (MMPDE5) using finite differences in space and the θ method in time, $\frac{1}{2} \leq \theta \leq 1$, to obtain the update from $t^{[\nu]}$ to $t^{[\nu+1]}$ as the solution of

$$\begin{aligned} \frac{x_i^{[\nu+1]} - x_i^{[\nu]}}{\Delta t} - \frac{\theta}{\tau \Delta \xi^2} \left(\tilde{M}(x_i^{[\nu+1]}, x_{i+1}^{[\nu+1]})(x_{i+1}^{[\nu+1]} - x_i^{[\nu+1]}) - \tilde{M}(x_{i-1}^{[\nu+1]}, x_i^{[\nu+1]})(x_i^{[\nu+1]} - x_{i-1}^{[\nu+1]}) \right) \\ - \frac{(1-\theta)}{\tau \Delta \xi^2} \left(\tilde{M}(x_i^{[\nu]}, x_{i+1}^{[\nu]})(x_{i+1}^{[\nu]} - x_i^{[\nu]}) - \tilde{M}(x_{i-1}^{[\nu]}, x_i^{[\nu]})(x_i^{[\nu]} - x_{i-1}^{[\nu]}) \right) &= 0, \\ x_0^{[\nu+1]} = x_L \quad x_N^{[\nu+1]} = x_R. \end{aligned} \quad (1.3)$$

for $i = 1, \dots, N-1$.

The goal of this paper is to prove that whenever the number of grid points N is large enough, the parallel and alternating Schwarz algorithms applied to the *discrete* problems (1.2) and (1.3) do indeed converge to the discrete monodomain solution. We will also estimate the rate of convergence in terms of the number of grid points and parameters related to the monitor function M . It is important to note that the analysis of the discrete algorithms is not a simple extension of the continuous theory presented in [22, 28]. That analysis depended heavily on implicit integral representations of the subdomain solutions and on properties of those integral representations. These properties do not carry over once quadrature has been applied. In fact, even in the single-domain case, the discretized problem may not be well-posed when the number of grid points is not large enough, see Example 2.1, and also Xu et al. [48], where the authors considered quadrature using piecewise constant or linear polynomials in the context of de Boor's algorithm. Instead, the theory here relies heavily on the notions of nonlinear M -functions and the method of super- and subsolutions in \mathbb{R}^n popularized by Ortega and Rheinboldt, see for example [44, 41].

The literature involving DD approaches for linear elliptic PDEs is vast, see the extensive books [47, 39]. We focus in this paper on Schwarz methods, since these methods have been studied for time dependent problems in two fundamentally different approaches: the more classical approach, in which one applies the alternating or parallel Schwarz method to the sequence of elliptic problems obtained from semi-discretization (in time) of the PDE [8, 9], or the more recent approach, where one decomposes the space-time domain and applying the Schwarz waveform relaxation algorithm [5, 24, 25]. For both approaches, optimized variants have been developed, see

[18] and the references therein for the first, and [21, 19, 20, 2] for the second approach. The linear theory, however, does not apply to the mesh PDEs considered here, since they are highly nonlinear. While DD methods have been applied successfully to a large variety of nonlinear problems (see e.g. [27, 3] for porous media flow problems, [45, 1] for problems from physiology, [26] for nonlinear parabolic systems, and [14] for the compressible Navier–Stokes equations), comparably much less has been written about the analysis of DD methods for nonlinear PDEs; see [10, 38, 37, 36, 16, 6, 46] for the steady case, and [17, 23] for evolution problems.

There are two common alternatives to the approach described in this paper. In the first, known as the Newton–Krylov–Schwarz (NKS) method, one discretizes the global nonlinear problem and solves the nonlinear algebraic equations with an inexact Newton method. The linear system that produces a Newton step is solved using a preconditioned Krylov method, where the inner (linear) iterations can be parallelized using a DD preconditioner. In the second approach, known as the nonlinear Additive Schwarz Preconditioned Inexact Newton (ASPIN) method, the discretized global nonlinear problem is not solved directly, but is replaced by a ‘preconditioned’ nonlinear function that has the same set of solutions as the original. Typically, the preconditioned function is a sum of local solutions of subdomain problems, so that it can be evaluated in parallel; thus, it is the analogue of the additive Schwarz method for linear problems. The preconditioned nonlinear problem is then solved using an inexact Newton method without additional linear preconditioning. See [11, 12, 42, 34] for more details. More recently, a two-level ASPIN algorithm has been proposed and studied in [35]. The multi-level approach and coarse space allows one to deal with deterioration of the convergence rate of the DD algorithm as the number of subdomains increases.

The paper is organized as follows. In Section 2, we begin by illustrating an example of (1.2) which is not well-posed. Then we introduce notation and review the main theoretical tools used in our analysis, namely the theory of M -functions. We then proceed in Sections 2.2 and 2.3 to specify sufficient conditions which guarantee the discrete systems for steady and time dependent mesh generation are M -functions. Our convergence results rely on the existence of super and subsolutions for the discrete systems; this is demonstrated in Section 3. Section 4 provides the main convergence results for both the parallel and alternating Schwarz methods for the discrete mesh generators. The asymptotic convergence rate of the discrete parallel Schwarz is given in Section 5, where we also show that the convergence rate coincides with the continuous analysis in the fine-grid limit. In Section 6, we illustrate our theoretical results with numerical examples, show the increased robustness of the DD approach relative to Newton’s method, and illustrate the use of DD within the MP iteration approach. Finally, we draw some conclusions in Section 7.

2. Uniqueness of solutions to the discrete equidistribution problem.

As we mentioned in the introduction, once the equidistribution equation (1.1) is discretized by a quadrature rule, it is no longer clear that a unique solution always exists. We start this section by an example showing non-uniqueness of the discrete solution when there are not enough points on the grid.

EXAMPLE 2.1. Consider the positive function $M : [0, 1] \rightarrow \mathbb{R}^+$ shown in Figure 2.1, defined as

$$M(x) = \begin{cases} 1 + 30x - \frac{171}{4}x^2, & 0 \leq x \leq \frac{2}{3}, \\ \frac{4-5x}{2x-1}, & \frac{2}{3} \leq x \leq \frac{3}{4}, \\ \frac{1}{7}(836 - 2496x + 2328x^2 - 640x^3), & \frac{3}{4} \leq x \leq 1. \end{cases}$$

This M is chosen so that $M(0) = 1$, $M(1) = 4$, and the function is continuously

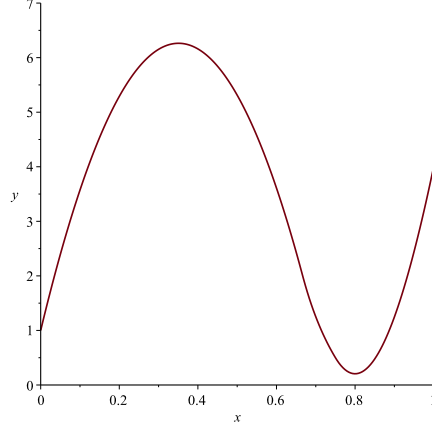


FIG. 2.1. An example mesh density function leading to a non-unique mesh transformation.

differentiable at $2/3$ and $3/4$. Suppose we wish to solve the steady mesh generation problem with exactly one interior node α , i.e., we want to find $\alpha \in (0, 1)$ such that

$$\int_0^\alpha M(u) du = \int_\alpha^1 M(u) du.$$

The solution to the continuous problem is unique, with $\alpha \approx 0.38$. However, if we replace the integral by the trapezoidal rule, i.e., if we solve

$$\alpha \frac{M(0) + M(\alpha)}{2} = (1 - \alpha) \frac{M(\alpha) + M(1)}{2}, \quad (2.1)$$

Then we observe that for $\frac{2}{3} \leq \alpha \leq \frac{3}{4}$, the left hand side of (2.1) becomes

$$\alpha \frac{M(0) + M(\alpha)}{2} = \frac{\alpha(2\alpha - 1 + 4 - 5\alpha)}{2(2\alpha - 1)} = \frac{3}{2} \cdot \frac{1 - \alpha}{2\alpha - 1},$$

and the right hand side gives

$$(1 - \alpha) \frac{M(\alpha) + M(1)}{2} = \frac{(1 - \alpha)4 - 5\alpha + 4(2\alpha - 1)}{2(2\alpha - 1)} = \frac{3}{2} \cdot \frac{1 - \alpha}{2\alpha - 1}.$$

Thus, every α in the interval $[\frac{2}{3}, \frac{3}{4}]$ is a solution of (2.1), so the discrete problem, defined on only 3 nodes, is not well-posed.

The problem in the above example is that there are too few points in the grid to “nail down” the area under the curve with sufficient precision to obtain a unique

solution. One of the main theoretical results of this paper states that whenever the grid is fine enough (where ‘fine enough’ depends on the mesh density function M), the discrete problem also has a unique solution. The main idea is to show that the equidistribution equations satisfy the M -function property in the sense of Ortega and Rheinboldt [44, 41] whenever the grid is fine enough. Uniqueness then follows from the M -function structure; in addition, the convergence of the parallel and alternating Schwarz methods is also a consequence of this M -function structure.

2.1. M -functions: definitions and properties. Consider a finite dimensional system of nonlinear equations $F(\mathbf{x}) = 0$ where $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is a vector of unknowns. We will refer to the i th equation in the system as $f_i(\mathbf{x})$. The natural partial ordering (component-wise) on \mathbb{R}^n is defined by

$$\mathbf{x} \leq \mathbf{y}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad \text{if and only if} \quad x_i \leq y_i, \text{ for } i = 1, \dots, n.$$

We will let $\mathbf{e}^i, i = 1, \dots, n$ denote the i -th standard basis vector in \mathbb{R}^n .

We begin by defining isotone and antitone mappings and then diagonally isotone and off-diagonally antitone mappings, see [41].

DEFINITION 2.2. *A mapping $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is isotone (antitone) if for all $\mathbf{x}, \mathbf{y} \in \mathbb{D}$, $F(\mathbf{x}) \leq F(\mathbf{y})$ ($F(\mathbf{x}) \geq F(\mathbf{y})$) whenever $\mathbf{x} \leq \mathbf{y}$. It is strictly isotone (antitone) if $F(\mathbf{x}) < F(\mathbf{y})$ ($F(\mathbf{x}) > F(\mathbf{y})$) whenever $\mathbf{x} < \mathbf{y}$.*

DEFINITION 2.3. *A mapping $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is diagonally isotone if, for any $\mathbf{x} \in \mathbb{D}$, the n functions $f_i(\mathbf{x} + t\mathbf{e}^i), i = 1, \dots, n$, are isotone when $\mathbf{x} + t\mathbf{e}^i \in \mathbb{D}$. F is off-diagonally antitone if, for any $\mathbf{x} \in \mathbb{R}^n$, the functions $f_i(\mathbf{x} + t\mathbf{e}^j), i \neq j, i, j = 1, \dots, n$, are antitone when $\mathbf{x} + t\mathbf{e}^i \in \mathbb{D}$.*

Collatz [13] introduced the following converse notion of isotonicity on partially ordered topological spaces.

DEFINITION 2.4. *A mapping $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is inverse isotone if for any $\mathbf{x}, \mathbf{y} \in \mathbb{D}$, $F(\mathbf{x}) \leq F(\mathbf{y})$ implies $\mathbf{x} \leq \mathbf{y}$.*

The following notion of an M -function was originally introduced by Ortega and developed by Rheinboldt [44].

DEFINITION 2.5. *A mapping $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an M -function if F is inverse isotone and off-diagonally antitone.*

THEOREM 2.6 (Uniqueness). *Let $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be an M -function and consider the problem $F(\mathbf{x}) = \mathbf{b}$, where $\mathbf{b} \in \mathbb{R}^n$ is a given right hand side. Then the solution $\mathbf{x} \in \mathbb{D}$, if it exists, must be unique.*

Proof. Let \mathbf{x} and \mathbf{y} be two solutions satisfying $F(\mathbf{x}) = F(\mathbf{y}) = \mathbf{b}$. Then since $F(\mathbf{x}) \leq F(\mathbf{y})$, the inverse isotonicity of F implies that $\mathbf{x} \leq \mathbf{y}$. On the other hand, we also have $F(\mathbf{y}) \leq F(\mathbf{x})$, which implies $\mathbf{y} \leq \mathbf{x}$. It follows that $\mathbf{x} = \mathbf{y}$, i.e., the solution is unique. \square

Of course, the difficulty in determining if a map is a M -function is checking the inverse isotonicity. A sufficient condition requires the following definition, which first appeared in [44] and specifies the connectivity of the unknowns through F .

DEFINITION 2.7. *Let $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ be off-diagonally antitone. Then a link (i, j) is strict if the function $t \mapsto f_i(\mathbf{x} + t\mathbf{e}^j)$ is strictly antitone. A path $i \rightsquigarrow j$ exists if there exists a sequence of strict links $(i, j_1), (j_1, j_2), \dots, (j_k, j)$.*

A sufficient condition for $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ to be an M -function is given in the following lemma, a modification of Theorem 5.1 in Rheinboldt [44]. See the Appendix for a proof of this result.

LEMMA 2.8. Let \mathbb{D} be a convex and open subset of \mathbb{R}^n . Assume $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is off-diagonally antitone, and that for any $\mathbf{x} \in \mathbb{D}$, the functions $q_i : S_i \subset \mathbb{R} \rightarrow \mathbb{R}^n$ defined as

$$q_i(\tau) = \sum_{j=1}^n f_j(\mathbf{x} + \tau \mathbf{e}^i), \quad i = 1, \dots, n, \quad \text{with} \quad S_i = \{\tau : \mathbf{x} + \tau \mathbf{e}^i \in \mathbb{D}\},$$

are isotone. Finally, assume for every i that either q_i is strictly isotone, or there exists a node $j = j(i)$ such that a path $j \rightsquigarrow i$ exists and q_j is strictly isotone. Then F is an M -function.

Here we are interested in mappings F defined by the simple finite difference discretizations (1.2) and (1.3). We set $h = 1/N$ and choose the uniform grid $\xi_i = ih, i = 0, \dots, N$, in the computational domain. Since the values x_0, x_1, \dots, x_N in (1.2) and (1.3) are meant to be samples of a continuous function $x(\xi)$ at ξ_i , we expect successive samples x_{i-1} and x_i to be close when the grid is fine enough. Thus, we will consider the nonlinear mappings defined by (1.2) and (1.3) on the subset D of \mathbb{R}^{N+1} given by

$$D = \left\{ \mathbf{x} \in \mathbb{R}^{N+1} : |x_i - x_{i-1}| < \frac{\tilde{m}}{B}, \text{ for } i = 1, \dots, N \right\}, \quad (2.2)$$

where $\tilde{m} > 0$ and $B > 0$ are constants that depend on the function $M(x)$, to be defined later. From (2.2) it is clear that D is open and convex, as required by Lemma 2.8.

2.2. Steady mesh generation. Throughout this paper, we assume the two-point mesh density function $\tilde{M}(\cdot, \cdot)$ satisfies the following conditions:

(C1) there exists constants \tilde{m} and \hat{m} such that

$$0 < \tilde{m} \leq \tilde{M}(u, v) \leq \hat{m} < \infty \quad \text{for all } u, v \in \mathbb{R},$$

and

(C2) $\tilde{M}(\cdot, \cdot)$ is continuously differentiable with respect to each variable, and there exists a positive constant $B < \infty$ such that

$$\left| \frac{\partial \tilde{M}}{\partial u}(u, v) \right| \leq B, \quad \left| \frac{\partial \tilde{M}}{\partial v}(u, v) \right| \leq B \quad \text{for all } u, v \in \mathbb{R}.$$

(C3) for every $u, v \in \mathbb{R}$ such that $u \leq v$, there exists $\zeta \in [u, v]$ such that

$$\tilde{M}(u, v) = M(\zeta).$$

The condition (C3) simply assures consistency between M and \tilde{M} . The constants \tilde{m} , \hat{m} and B can typically be derived from the bounds and Lipschitz constant of the continuous mesh density function $M(x)$. For instance, for the trapezoidal rule $\tilde{M}(u, v) = \frac{1}{2}(M(u) + M(v))$, we have $B = \frac{1}{2} \max_{u \in \mathbb{R}} |M'(u)|$.

REMARK. Typically $M(x)$ is only defined on the interval $[0, 1]$, since the mesh is only defined on that interval. However, for the purpose of proving existence, we will need to evaluate M at points outside the unit interval; this can be done by extending $M(x)$ smoothly outside $[0, 1]$ in such a way that M becomes constant outside a slightly larger interval $[-x_{\min}, x_{\max}] \supset [0, 1]$. Thus, it is reasonable to assume that \tilde{m} and \hat{m} exist, and one can expect these bounds to be close to the maximum and minimum of

M inside the unit interval. A similar argument goes for the constant B , as long as the extension does not introduce gradients much steeper than those inside the unit interval.

We now use Lemma 2.8 to show that H defined by (1.2) is an M -function on D .

THEOREM 2.9. *The mapping $H : D \subset R^{N+1} \rightarrow R^{N+1}$ defined by*

$$h_i = \begin{cases} (\hat{m} + \check{m})(x_0 - x_L), & i = 0, \\ \tilde{M}(x_{i-1}, x_i)(x_i - x_{i-1}) - \tilde{M}(x_i, x_{i+1})(x_{i+1} - x_i), & i = 1, \dots, N-1, \\ (\hat{m} + \check{m})(x_N - x_R), & i = N, \end{cases} \quad (2.3)$$

is an M -function on the set D defined in (2.2) if $M(x)$ satisfies the conditions (C1) and (C2).

Proof. For $1 \leq i \leq N-1$, the i th component of H reads

$$h_i = \phi_{i-1/2}(x_{i-1}, x_i) - \phi_{i+1/2}(x_i, x_{i+1}),$$

where we have introduced the notation $\phi_{i-1/2} := \tilde{M}(x_{i-1}, x_i)(x_i - x_{i-1})$ representing the *discrete flux* between x_i and x_{i-1} .

The function h_i depends on x_j only for $j = i-1, i$ and $i+1$. Computing directly for $i = 1, \dots, N-1$ we have

$$\begin{aligned} \frac{\partial h_i}{\partial x_{i-1}} &= \frac{\partial \phi_{i-1/2}}{\partial x_{i-1}} = -\tilde{M}(x_{i-1}, x_i) + \tilde{M}_u(x_{i-1}, x_i)(x_i - x_{i-1}), \\ \frac{\partial h_i}{\partial x_{i+1}} &= -\frac{\partial \phi_{i+1/2}}{\partial x_{i+1}} = -\tilde{M}(x_i, x_{i+1}) - \tilde{M}_v(x_i, x_{i+1})(x_{i+1} - x_i). \end{aligned}$$

Here \tilde{M}_u and \tilde{M}_v denote the derivatives of $\tilde{M}(u, v)$ with respect to its first and second arguments respectively. Following Lemma 2.8, we begin by showing that F is off-diagonally antitone on the set D defined in (2.2).

Since the mesh $\{x_i\}_{i=0}^N$ belongs to D and \tilde{M} is Lipschitz with constant B (from condition (C2)), we have

$$\tilde{M}_u(x_{i-1}, x_i)(x_i - x_{i-1}) \leq |\tilde{M}_u(x_{i-1}, x_i)| \cdot |x_i - x_{i-1}| < B \frac{\check{m}}{B} = \check{m} \leq \tilde{M}(x_{i-1}, x_i)$$

and

$$-\tilde{M}_v(x_i, x_{i+1})(x_{i+1} - x_i) \leq |\tilde{M}_v(x_i, x_{i+1})| \cdot |x_{i+1} - x_i| < B \frac{\check{m}}{B} = \check{m} \leq \tilde{M}(x_i, x_{i+1}).$$

Hence $\frac{\partial h_i}{\partial x_{i-1}}$ and $\frac{\partial h_i}{\partial x_{i+1}}$ are strictly negative for $i = 1, \dots, N-1$. Furthermore, $\partial h_0 / \partial x_i = 0$ for $i = 1, \dots, N$ and $\partial h_N / \partial x_i = 0$ for $i = 0, \dots, N-1$ which implies H is off-diagonally antitone.

We now construct the functions $q_i(t) = \sum_{j=0}^N h_j(\mathbf{x} + t\mathbf{e}^i)$, $i = 1, \dots, N-1$, and take derivatives with respect to t . Direct calculation gives $dq_i/dt = 0$ for $i = 2, \dots, N-2$, while

$$\begin{aligned} \frac{dq_0}{dt} &= \hat{m} + \check{m} - \tilde{M}(x_0 + t, x_1) + \tilde{M}_u(x_0 + t, x_1)(x_1 - x_0 - t), \\ \frac{dq_1}{dt} &= \frac{d}{dt} \phi_{1/2}(x_0, x_1 + t) = \tilde{M}_v(x_0, x_1 + t)(x_1 + t - x_0) + \tilde{M}(x_0, x_1 + t), \end{aligned}$$

and

$$\begin{aligned}\frac{dq_{N-1}}{dt} &= -\tilde{M}_u(x_{N-1} + t, x_N)(x_N - x_{N-1} - t) + \tilde{M}(x_{N-1} + t, x_N), \\ \frac{dq_N}{dt} &= \hat{m} + \check{m} - \tilde{M}(x_{N-1}, x_N + t) - \tilde{M}_v(x_{N-1}, x_N + t)(x_N + t - x_{N-1}).\end{aligned}$$

Using condition (C2) and the assumption that the perturbed mesh $x + te^i$ belongs to D it is easy to see that $\frac{dq_0}{dt}$, $\frac{dq_1}{dt}$, $\frac{dq_{N-1}}{dt}$, $\frac{dq_N}{dt}$ are strictly positive. Hence the functions q_i are isotone. Finally, the nodes x_1 and x_{N-1} are connected to all other nodes, with q_1 and q_{N-1} strictly isotone. So by Lemma 2.8, the function H defined by (2.3) is an M -function on D . \square

2.3. Time dependent mesh generation. Given a physical solution at time level $t^{[\nu]}$, the updated mesh at time $t^{[\nu+1]}$ can be found by solving the nonlinear system (1.3). We now verify that this system is described by a M -function.

THEOREM 2.10. *The mapping $G : D \subset \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$, for the unknowns $\{x_i^{[\nu+1]}\}_{i=0}^N$, defining (1.3) is an M -function on the set D defined in (2.2) if $\tilde{M}(u, v)$ satisfies the conditions (C1) and (C2).*

Proof. For ν fixed, the unknowns $x^{[\nu+1]}$ satisfy $G(x^{[\nu+1]}) = 0$ where from (1.3) the i th component of G reads

$$g_i = \begin{cases} \theta\gamma(\hat{m} + \check{m})(x_0^{[\nu+1]} - x_L), & i = 0 \\ x_i^{[\nu+1]} - \theta\gamma(\phi_{i+1/2}^{[\nu+1]} - \phi_{i-1/2}^{[\nu+1]}) - x_i^{[\nu]} - (1 - \theta)\gamma(\phi_{i+1/2}^{[\nu]} - \phi_{i-1/2}^{[\nu]}), & i = 1, \dots, N-1, \\ \theta\gamma(\hat{m} + \check{m})(x_N^{[\nu+1]} - x_R), & i = N \end{cases} \quad (2.4)$$

where we have introduced the notation $\phi_{i-1/2}^{[\nu]} := \tilde{M}(x_{i-1}^{[\nu]}, x_i^{[\nu]})(x_i^{[\nu]} - x_{i-1}^{[\nu]})$ and $\gamma := \Delta t / (\tau \Delta \xi^2)$.

Computing directly for $i = 1, \dots, N-1$ we have

$$\begin{aligned}\frac{\partial g_i}{\partial x_{i-1}^{[\nu+1]}} &= \theta\gamma \frac{\partial \phi_{i-1/2}^{[\nu+1]}}{\partial x_{i-1}^{[\nu+1]}} = \theta\gamma(-\tilde{M}(x_{i-1}^{[\nu+1]}, x_i^{[\nu+1]}) + \tilde{M}_u(x_{i-1}^{[\nu+1]}, x_i^{[\nu+1]})(x_i^{[\nu+1]} - x_{i-1}^{[\nu+1]})), \\ \frac{\partial g_i}{\partial x_{i+1}^{[\nu+1]}} &= -\theta\gamma \frac{\partial \phi_{i+1/2}^{[\nu+1]}}{\partial x_{i+1}^{[\nu+1]}} = -\theta\gamma(\tilde{M}(x_i^{[\nu+1]}, x_{i+1}^{[\nu+1]}) + \tilde{M}_v(x_i^{[\nu+1]}, x_{i+1}^{[\nu+1]})(x_{i+1}^{[\nu+1]} - x_i^{[\nu+1]})).\end{aligned}$$

Once again \tilde{M}_u and \tilde{M}_v denote the derivatives of $\tilde{M}(u, v)$ with respect to its first and second arguments respectively. Moreover $\partial g_0 / \partial x_i = 0$ for $i = 1, \dots, N$ and $\partial g_N / \partial x_i = 0$ for $i = 0, \dots, N-1$. Arguing exactly as in the proof of Theorem 2.9 we can see that G is off-diagonally antitone on the set D defined in (2.2).

We now construct the functions $q_i(t)$ specified in Lemma 2.8. and take derivatives with respect to t . Direct calculation gives $dq_i/dt = 1$ (strictly positive) for $i = 2, \dots, N-2$. Also, we have

$$\begin{aligned}\frac{dq_0}{dt} &= \theta\gamma(\hat{m} + \check{m}) + \theta\gamma(\tilde{M}_u(x_0^{[\nu+1]} + t, x_1^{[\nu+1]})(x_1^{[\nu+1]} - x_0^{[\nu+1]} - t) - \tilde{M}(x_0^{[\nu+1]} + t, x_1^{[\nu+1]})), \\ \frac{dq_1}{dt} &= 1 + \theta\gamma(\tilde{M}_v(x_0^{[\nu+1]}, x_1^{[\nu+1]} + t)(x_1^{[\nu+1]} + t - x_0^{[\nu+1]}) + \tilde{M}(x_0^{[\nu+1]}, x_1^{[\nu+1]} + t)),\end{aligned}$$

and

$$\begin{aligned}\frac{dq_{N-1}}{dt} &= 1 - \theta\gamma(\tilde{M}_u(x_{N-1}^{[\nu+1]} + t, x_N^{[\nu+1]})(x_N^{[\nu+1]} - x_{N-1}^{[\nu+1]} - t) - \tilde{M}(x_{N-1}^{[\nu+1]} + t, x_N^{[\nu+1]})), \\ \frac{dq_N}{dt} &= \theta\gamma(\hat{m} + \check{m}) - \theta\gamma(\tilde{M}_v(x_{N-1}^{[\nu+1]}, x_N^{[\nu+1]})(x_N^{[\nu+1]} + t - x_{N-1}^{[\nu+1]}) + \tilde{M}(x_{N-1}^{[\nu+1]}, x_N^{[\nu+1]})).\end{aligned}$$

Using condition (C2) and the assumption that the perturbed mesh $x^{[\nu+1]} + te^i$ belongs to D it is easy to see that $\frac{dq_0}{dt}$, $\frac{dq_1}{dt}$, $\frac{dq_{N-1}}{dt}$, $\frac{dq_N}{dt}$ are strictly positive. Hence the functions q_i are strictly isotone for all i , so by Lemma 2.8 the function G defined by 2.4 is a M -function. \square

It follows from Theorems 2.9 and 2.10 that the steady and time-dependent mesh generation problems have unique solutions whenever the grid is fine enough. The existence of solutions, though, must be proved separately. In classical M -function theory, one proves the existence of solutions to $F(\mathbf{x}) = 0$ as follows [41]: First, one needs to construct a super- and a subsolution, i.e., vectors $\hat{\mathbf{x}}$ and $\check{\mathbf{x}}$ such that $\check{\mathbf{x}} \leq \hat{\mathbf{x}}$ and $F(\check{\mathbf{x}}) \leq 0 \leq F(\hat{\mathbf{x}})$. Next, one shows that the *nonlinear* Gauss–Seidel and Jacobi processes, starting with $\hat{\mathbf{x}}$ or $\check{\mathbf{x}}$, converge monotonically to a fixed point \mathbf{x}^* satisfying $F(\mathbf{x}^*) = 0$. Thus, the fixed point is a solution of the nonlinear problem and, by uniqueness, its only solution.

In the next section, we will construct super- and subsolutions to the steady and time-dependent mesh generation problems for grids that are fine enough. The existence of solutions then follows from the proof of convergence of a Jacobi or Gauss–Seidel process. This proof of convergence will be omitted, since it is identical to the one presented in [41]; however, in Section 4, we will show convergence of the parallel and alternating Schwarz methods, which are essentially block variants of the Jacobi and Gauss–Seidel processes.

3. Existence of Solutions. The existence of solutions as well as the convergence of our domain decomposition methods depend on the existence of a supersolution and a subsolution of the initial guess \mathbf{x}^0 , both of which must lie within the domain D of F . In both the steady and time-dependent mesh generation problems, the functions

$$h_i(u_{i-1}, u_i, u_{i+1}) := \tilde{M}(u_{i-1}, u_i)(u_i - u_{i-1}) - \tilde{M}(u_i, u_{i+1})(u_{i+1} - u_i),$$

for $i = 1, \dots, N$, already defined in (2.3), feature prominently in the nonlinear systems. To prove the existence of super- and subsolutions, we need to know which values can be attained by h_i when the u_i are not allowed to be too far apart. The following technical lemma provides an answer to this question.

LEMMA 3.1. *Let $N > B\hat{m}/\check{m}^2$, where B , \hat{m} and \check{m} are defined in conditions (C1) and (C2). If $c \in \mathbb{R}$ is such that*

$$|u_R - u_L| \frac{\hat{m}}{N} + N|c| < \frac{\check{m}^2}{B}, \quad (3.1)$$

then there exists a sequence $\{u_i\}_{i=0}^N$ such that $u_0 = u_L$, $u_N = u_R$, and

$$h_i(u_{i-1}, u_i, u_{i+1}) = c$$

for $i = 1, 2, \dots, N - 1$. Moreover, the u_i satisfy

$$|u_i - u_{i-1}| \leq \frac{1}{\check{m}} \left(|u_R - u_L| \frac{\hat{m}}{N} + N|c| \right) < \frac{\check{m}}{B}.$$

Proof. Let us consider the flux $\phi_{i-1/2}$ between grid points $i-1$ and i given by

$$\phi_{i-1/2} = \tilde{M}(u_{i-1}, u_i)(u_i - u_{i-1}).$$

Using this notation, we see that $h_i(\mathbf{u}) = \phi_{i+1/2} - \phi_{i-1/2}$, so if we fix a parameter $s \in \mathbb{R}$ and require the fluxes $\phi_{i-1/2}$ to satisfy

$$\phi_{i-1/2} = \phi_{i-1/2}(s) = s + \sigma\chi_{i-1/2}, \quad (3.2)$$

where $\chi_{i-1/2} := 1 - \frac{2i-1}{N}$ is a decreasing sequence between -1 and 1 , then

$$h_i(\mathbf{u}(s)) = \phi_{i-1/2}(s) - \phi_{i+1/2}(s) = \sigma(\chi_{i-1/2} - \chi_{i+1/2}) = \frac{2\sigma}{N}.$$

So if we choose $\sigma = Nc/2$, then $h_i = c$ for all i . The extra parameter s , despite not having an influence on the value of h_i , is important because it will eventually allow us to satisfy the boundary conditions $u_0 = u_L$, $u_N = u_R$. (This is similar in spirit to shooting methods for solving two-point boundary value problems; here, s plays the role of the initial velocity, which will be chosen to satisfy the boundary condition $u_N = u_R$.)

Our next task is to show that for any parameter value s satisfying $|s| < \tilde{m}/B - |\sigma|$, we can indeed choose the fluxes as specified in (3.2), i.e., the recurrence $u_0(s) = u_L$,

$$\tilde{M}(u_{i-1}(s), u_i(s))(u_i(s) - u_{i-1}(s)) = \phi_{i-1/2}(s), \quad i = 1, \dots, N \quad (3.3)$$

is solvable for these values of s and the resulting \mathbf{u} is in the set D . To show that this is possible, consider two points $t < \tilde{t}$, with $\max\{|t - u_{i-1}|, |\tilde{t} - u_{i-1}|\} < \tilde{m}/B$. Then

$$\begin{aligned} & \tilde{M}(u_{i-1}, \tilde{t})(\tilde{t} - u_{i-1}) - \tilde{M}(u_{i-1}, t)(t - u_{i-1}) \\ &= \tilde{M}(u_{i-1}, \tilde{t})(\tilde{t} - t) + [\tilde{M}(u_{i-1}, \tilde{t}) - \tilde{M}(u_{i-1}, t)](t - u_{i-1}) \\ &\geq \tilde{m}(\tilde{t} - t) - B(\tilde{t} - t)|t - u_{i-1}| > 0. \end{aligned}$$

Thus, the function $t \mapsto \tilde{M}(u_{i-1}, t)(t - u_{i-1})$ is strictly increasing within the interval $\{t : |t - u_{i-1}| < \tilde{m}/B\}$, and since

$$|\tilde{M}(u_{i-1}, t)(t - u_{i-1})| \geq \tilde{m}|t - u_{i-1}|,$$

its image contains the interval $(-\tilde{m}^2/B, \tilde{m}^2/B)$. Thus, for

$$|s| < \tilde{m}^2/B - |\sigma|,$$

we have $|\phi_{i-1/2}(s)| < \tilde{m}^2/B$ for all i , so (3.3) is uniquely solvable for $u_i(s)$.

Finally, we show that the boundary condition $u_N = u_R$ can be satisfied by choosing the parameter s appropriately. We have

$$u_N(s) - u_L = \sum_{i=1}^N (u_i(s) - u_{i-1}(s)) = \sum_{i=1}^N \frac{\phi_{i-1/2}(s)}{\tilde{M}(u_{i-1}(s), u_i(s))}.$$

Substituting the relation (3.2) into $\phi_{i-1/2}(s)$ then gives

$$u_N(s) - u_L = \sum_{i=1}^N \frac{s + \sigma\chi_{i-1/2}}{\tilde{M}(u_{i-1}(s), u_i(s))}.$$

Let $\alpha = (u_R - u_L)/N$. If $\alpha \geq 0$, then letting $s_+ = \alpha\hat{m} + |\sigma|$ leads to

$$u_N(s_+) \geq u_L + \sum_{i=1}^N \frac{(u_R - u_L)\hat{m}/N}{\tilde{M}(u_{i-1}(s), u_i(s))} \geq u_L + u_R - u_L = u_R,$$

and letting $s_- = \alpha\check{m} - |\sigma|$ leads to

$$u_N(s_-) \leq u_L + \sum_{i=1}^N \frac{(u_R - u_L)\check{m}/N}{\tilde{M}(u_{i-1}(s), u_i(s))} \leq u_L + u_R - u_L = u_R.$$

If $\alpha < 0$ (i.e., if $u_L > u_R$), then a similar calculation holds by letting $s_+ = \alpha\check{m} + |\sigma|$ and $s_- = \alpha\hat{m} - |\sigma|$. Since u_N is a continuous function of s , there exists s^* between s_- and s_+ such that $u_N(s^*) = u_R$, as long as $|s_{\pm}| < \check{m}^2/B - |\sigma|$. A sufficient condition is

$$|u_R - u_L| \frac{\hat{m}}{N} + N|c| < \frac{\check{m}^2}{B},$$

since this implies

$$|s_{\pm}| \leq |u_R - u_L| \frac{\hat{m}}{N} + \frac{N|c|}{2} < \frac{\check{m}^2}{B} - \frac{N|c|}{2} = \frac{\check{m}^2}{B} - |\sigma|.$$

Finally, we have

$$|u_i - u_{i-1}| = \frac{|\phi_{i-1/2}|}{\tilde{M}(u_{i-1}, u_i)} \leq \frac{s_+ + |\sigma|}{\check{m}} < \frac{\check{m}}{B},$$

so $\mathbf{u} \in D$, i.e., the solution stays in the region where h is an M -function. \square

Lemma 3.1 immediately implies the existence of solutions to the steady-state problem when the grid is fine enough.

THEOREM 3.2. *Let $N > B\hat{m}/\check{m}^2$. Then there exists a vector $\mathbf{x} \in D$ that satisfies the steady-state equidistribution problem $H(\mathbf{x}) = 0$ defined in (1.2).*

Proof. Letting $c = 0$, $u_L = 0$ and $u_R = 1$ in Lemma 3.1 gives the required solution, with $x_i = u_i$ for $i = 0, 1, \dots, N$. \square

For the time-dependent case, we can reformulate (1.3) as

$$\begin{aligned} g_i(u_{i-1}, u_i, u_{i+1}) &= (u_i - x_i^{[\nu]}) - \theta\gamma(\phi_{i+1/2}(u_i, u_{i+1}) - \phi_{i-1/2}(u_{i-1}, u_i)) \\ &\quad - (1 - \theta)\gamma(\phi_{i+1/2}(x_i^{[\nu]}, x_{i+1}^{[\nu]}) - \phi_{i-1/2}(x_{i-1}^{[\nu]}, x_i^{[\nu]})) = 0, \end{aligned} \tag{3.4}$$

where $\gamma = \Delta t/(\tau\Delta\xi^2)$. The above equation is to be solved for $u_i = x_i^{[\nu+1]}$, $i = 1, \dots, N-1$.

To guarantee the existence of a solution to this problem, we need to make an assumption on $\mathbf{x}^{[\nu]}$, the solution at the previous time step. We assume that there exists a constant κ (the ‘‘curvature’’) such that the residuals $r_i := g_i(x_{i-1}^{[\nu]}, x_i^{[\nu]}, x_{i+1}^{[\nu]}) = h_i(x_{i-1}^{[\nu]}, x_i^{[\nu]}, x_{i+1}^{[\nu]})$ satisfy

$$|r_i| \leq \kappa/N^2. \tag{3.5}$$

This is true if, for instance, $x_i^{[\nu]}$ are equidistant samples of a smooth function $x^{[\nu]}(\xi)$: in this case, we can write $r_i = \gamma(\psi(1) - \psi(0)) = \gamma\psi'(\eta)$ with $\eta \in (0, 1)$, where

$$\psi(t) = \tilde{M}(tx_{i-1}^{[\nu]} + (1-t)x_i^{[\nu]}, tx_i^{[\nu]} + (1-t)x_{i+1}^{[\nu]})[t(x_i^{[\nu]} - x_{i-1}^{[\nu]}) + (1-t)(x_{i+1}^{[\nu]} - x_i^{[\nu]})].$$

Computing directly we have

$$\begin{aligned} \psi'(\eta) &= [\tilde{M}_u(\cdot)(x_{i-1}^{[\nu]} - x_i^{[\nu]}) + \tilde{M}_v(\cdot)(x_i^{[\nu]} - x_{i+1}^{[\nu]})][\eta(x_i^{[\nu]} - x_{i-1}^{[\nu]}) + (1-\eta)(x_{i+1}^{[\nu]} - x_i^{[\nu]})] \\ &\quad + M(\cdot)[2x_i^{[\nu]} - x_{i-1}^{[\nu]} - x_{i+1}^{[\nu]}]. \end{aligned}$$

If $|x_\xi| < L_x$ and $|x_{\xi\xi}| < K_x$, then

$$|\psi'(\eta)| \leq \frac{1}{N^2}(2BL_x^2 + \hat{m}K_x).$$

That is, (3.5) holds with

$$\kappa = \gamma(2BL_x^2 + \hat{m}K_x).$$

THEOREM 3.3. *Let $N \geq B(\hat{m} + \kappa)/\check{m}^2$ where B , \hat{m} , and \check{m} are defined in conditions (C1) and (C2) and κ is defined in (3.5). Then for $\frac{1}{2} \leq \theta \leq 1$, there exists a unique solution $\{u_i\}_{i=0}^N$ to (3.4).*

Proof. Assuming (3.5) holds, we can construct a supersolution for the time-dependent problem. Let $c = \kappa/N^2$, $u_0 = 0$, $u_N = 1$, and $\hat{\mathbf{u}}$ be the vector constructed from this value using Lemma 3.1. For this to work, we need

$$\frac{\check{m}^2}{B} > \frac{\hat{m}}{N} + N|c| = \frac{\hat{m} + \kappa}{N},$$

or

$$N > \frac{B(\hat{m} + \kappa)}{\check{m}^2}.$$

Then we have

$$h_i(\hat{\mathbf{u}}) = c = \frac{\kappa}{N^2} \geq r_i = g_i(\mathbf{x}^{[\nu]}) = h_i(\mathbf{x}^{[\nu]}).$$

Since H is an M -function, this implies $\hat{u}_i \geq x_i^{[\nu]}$ for all i . Thus, we have

$$g_i(\hat{\mathbf{u}}) \geq \theta h_i(\hat{\mathbf{u}}) + (1-\theta)r_i \geq (1-\theta)(\kappa/N^2 - |r_i|) \geq 0,$$

where we have used the fact that $\theta \geq 1-\theta$, since $\theta \in [\frac{1}{2}, 1]$. Thus, $\hat{\mathbf{u}}$ is a supersolution. Similarly, we can choose $c = -\kappa/N^2$ to construct a subsolution $\check{\mathbf{u}}$, with the same condition on N as above. Now, using $\hat{\mathbf{u}}$ or $\check{\mathbf{u}}$ as the initial guess, one can show [41, §13.5] that the nonlinear Jacobi/Gauss-Seidel process converges to a vector \mathbf{u} satisfying $G(\mathbf{u}) = 0$, so the solution exists (and is unique by Theorem 2.6). \square

REMARK. Since $g_i(\mathbf{x}^{[\nu]}) = h_i(\mathbf{x}^{[\nu]}) = r_i$, we in fact have

$$g_i(\hat{\mathbf{u}}) \geq \theta h_i(\hat{\mathbf{u}}) + (1-\theta)r_i \geq \theta h_i(\mathbf{x}^{[\nu]}) + (1-\theta)r_i = r_i,$$

which implies $g_i(\hat{\mathbf{u}}) \geq g_i(\mathbf{x}^{[\nu]})$. Similarly, we have $g_i(\check{\mathbf{u}}) \leq g_i(\mathbf{x}^{[\nu]})$. These inequalities will be used in the next section to show that the alternating and parallel Schwarz methods converge when $\mathbf{x}^{[\nu]}$ is used as an initial guess.

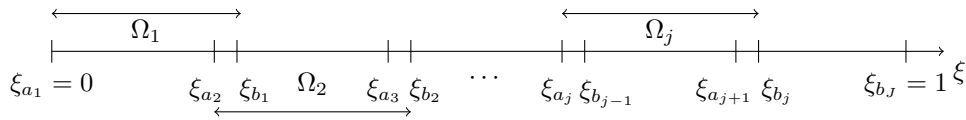


FIG. 4.1. *Decomposition into many subdomains.*

4. Convergence of Domain Decomposition Methods. In this section, we consider domain decomposition algorithms for solving the equidistribution problems (1.2) and (1.3). Each system may be written as $F(\mathbf{x}) = 0$, where $F : D \subset \mathbb{R}^{N+1} \rightarrow \mathbb{R}^{N+1}$ is given by $F = (f_0, \dots, f_N)^T$, where $f_i = h_i$ from (2.3) in the steady case and $f_i = g_i$ from (2.4) in the time dependent case. We have seen in Theorems 2.9 and 2.10 that in each case F is an M -function whenever the distance between successive points is not too large.

We now decompose the set of nodes $\{0, \dots, N\}$ into J overlapping subdomains as shown in Figure 4.1, i.e., we have $\Omega_j = \{i : a_j < i < b_j\}$, with $\xi_{a_1} = 0$, $\xi_{b_J} = 1$ and

$$a_j < b_{j-1} < a_{j+1} < b_j \quad \text{for } 2 \leq j \leq J-1.$$

The subdomain problem associated with Ω_j is now defined as follows: we solve $F_j(\mathbf{x}^j; x_{a_j}, x_{b_j}) = 0$ for \mathbf{x}^j , where

$$F_j = (f_{a_j+1}, \dots, f_{b_j-1})^T, \quad \mathbf{x}^j = (x_{a_j+1}, \dots, x_{b_j-1}^T), \quad (4.1)$$

and the boundary values at a_j and b_j are given by x_{a_j} and x_{b_j} . Note that since F_j is a restriction of F , the functions $F_j(\cdot; x_{a_j}, x_{b_j}) : \mathbb{R}^{b_j-a_j-1} \rightarrow \mathbb{R}^{b_j-a_j-1}$ are M -functions. Moreover, the subdomain functions F_j inherit the antitonicity property with respect to the boundary values, i.e., we have

$$F_j(\mathbf{x}^j; \alpha, \beta) \geq F_j(\mathbf{x}^j; \alpha', \beta') \quad (4.2)$$

whenever $\alpha \leq \alpha'$ and $\beta \leq \beta'$.

We can now define the classical alternating Schwarz (or Gauss–Seidel) algorithm: for $k = 0, 1, 2, \dots$ and for $j = 1, \dots, J$ solve

$$F_j(\mathbf{x}^{j,k+1}; x_{a_j}^{j-1,k+1}, x_{b_j}^{j+1,k}) = 0 \quad \text{for } \mathbf{x}^{j,k+1}. \quad (4.3)$$

Similarly, we can define the classical parallel Schwarz (or Jacobi) algorithm: for $k = 0, 1, 2, \dots$ and for $j = 1, \dots, J$, solve

$$F_j(\mathbf{x}^{j,k+1}; x_{a_j}^{j-1,k}, x_{b_j}^{j+1,k}) = 0 \quad \text{for } \mathbf{x}^{j,k+1}. \quad (4.4)$$

Under the assumption of the existence of an appropriate subsolution and supersolution, the goal of this section is to prove that both iterations converge to the same unique global solution of (4.1). Thus, these DD iterations can be used in place of Newton’s method, especially when the latter has trouble converging.

4.1. Convergence of the Alternating Schwarz iteration. We first prove the convergence of the alternating Schwarz iteration. We will show this by adapting the technique of [44] to handle the subdomain solves.

THEOREM 4.1. *Let N be large enough so that the global problem possesses a unique solution $\mathbf{x}^* \in D \subset \mathbb{R}^{N+1}$. Let $\mathbf{x}^0 \in D \subset \mathbb{R}^{N+1}$ be a supersolution, i.e.,*

$F(\mathbf{x}^0) \geq 0 = F(\mathbf{x}^*)$. Then the subdomain iterates $\mathbf{x}^{j,k}$ obtained from the classical alternating Schwarz method (4.3) with initial guess $\mathbf{x}^{j,0} = \mathbf{x}^0|_{\Omega_j}$ are well defined and satisfy

$$\mathbf{x}^{j,k} \geq \mathbf{x}^{j,k+1} \geq \mathbf{x}^*|_{\Omega_j}$$

for all $k \geq 0$ and $\mathbf{x}^{j,k} \rightarrow \mathbf{x}^*|_{\Omega_j}$ as $k \rightarrow \infty$; that is we have convergence from above. Similarly, if \mathbf{x}^0 is a subsolution, then $\mathbf{x}^{j,k} \uparrow \mathbf{x}^*|_{\Omega_j}$ as $k \rightarrow \infty$; that is we have convergence from below.

Proof. We only prove the supersolution case here, the subsolution case being similar. For any $1 \leq j \leq J$ and $k \geq 0$, we define the fractional iterate $\mathbf{x}^{k+j/J} \in \mathbb{R}^{N+1}$ as

$$\mathbf{x}_i^{k+j/J} = \begin{cases} x_i^{\ell,k+1}, & a_\ell \leq i < a_{\ell+1}, \quad \ell < j, \\ x_i^{j,k+1}, & a_j \leq i < b_j, \\ x_i^{\ell,k}, & \max\{b_j, a_\ell\} \leq i < a_{\ell+1}, \quad \ell > j. \end{cases} \quad (4.5)$$

In other words, each successive fractional iterate is obtained by replacing, after each subdomain solve, the portion of the global iterate corresponding to Ω_j by $\mathbf{x}^{j,k+1}$. We note that $\mathbf{x}^{k+j/J}$ is well defined, i.e., the subdomain problem (4.3) has a unique solution on Ω_j , whenever $\mathbf{x}^{k+(j-1)/J}$ is a supersolution in D , since F_j is an M -function on D and the restriction of $\mathbf{x}^{k+(j-1)/J}$ onto Ω_j is a supersolution with respect to F_j . Moreover, the solution of the subdomain problem satisfies $|x_i^{j,k+1} - x_{i-1}^{j,k}| < \check{m}/B$, so that $\mathbf{x}^{k+j/J} \in D$.

Thus, it suffices to show inductively for all $1 \leq j \leq J$ and $k \geq 0$ that $F(\mathbf{x}^{k+j/J}) \geq 0$ and

$$\mathbf{x}^{k+(j-1)/J} \geq \mathbf{x}^{k+j/J} \geq \mathbf{x}^*;$$

the conclusion then follows by restricting $\mathbf{x}^{k+j/J}$ to Ω_j . The assumption that \mathbf{x}^0 is a supersolution means that the base case $k = 0$ and $j = 0$ is verified, so assume inductively that $F(\mathbf{x}^{k+(j-1)/J}) \geq 0$ for some $1 \leq j \leq J$ and $k \geq 0$. By the induction hypothesis, we have

$$F_j(\mathbf{x}^{k+(j-1)/J}|_{\Omega_j}; x_{a_j}^{k+1}, x_{b_j}^k) \geq 0,$$

whereas by construction, we have

$$F_j(\mathbf{x}^{k+j/J}|_{\Omega_j}; x_{a_j}^{k+1}, x_{b_j}^k) = 0.$$

Since $F_j(\cdot; x_{a_j}^{k+1}, x_{b_j}^k)$ is an M -function, it is inverse isotone, which implies $\mathbf{x}^{k+(j-1)/J}|_{\Omega_j} \geq \mathbf{x}^{k+j/J}|_{\Omega_j}$. But $\mathbf{x}^{k+(j-1)/J}$ and $\mathbf{x}^{k+j/J}$ are identical everywhere outside Ω_j ; so $\mathbf{x}^{k+(j-1)/J} \geq \mathbf{x}^{k+j/J}$. Moreover, we have

$$f_i(\mathbf{x}^{k+j/J}) = \begin{cases} f_i(\mathbf{x}^{k+(j-1)/J}) \geq 0 & \text{for } i < a_j \text{ and for } i > b_j, \\ 0 & \text{for } a_j < i < b_j, \end{cases}$$

so it remains to show that $f_i(\mathbf{x}^{k+j/J}) \geq 0$ for $i \in \{a_j, b_j\}$. But since $x_i^{k+(j-1)/J} = x_i^{k+j/J}$ for $i \in \{a_j, b_j\}$ and $\mathbf{x}^{k+(j-1)/J} \geq \mathbf{x}^{k+j/J}$, the off-diagonal antitonicity of F implies

$$0 \leq f_i(\mathbf{x}^{k+(j-1)/J}) \leq f_i(\mathbf{x}^{k+j/J}), \quad i \in \{a_j, b_j\}.$$

So $F(\mathbf{x}^{k+j/J}) \geq 0$, as required. Thus, we have shown by induction $\{\mathbf{x}^{k+j/J}\}_{j,k}$ is a decreasing sequence bounded below by \mathbf{x}^* , so it converges to some global solution $\tilde{\mathbf{x}}^*$. But for any node $0 \leq i \leq N+1$, there is always a subsequence $\{r_m^{(i)}\}_{m \geq 0}$ for which $f_i(\mathbf{x}^{r_m^{(i)}}) = 0$ for all $m \geq 0$, so the continuity of F implies $F(\tilde{\mathbf{x}}^*) = 0$. Since F is an M -function, this implies $\tilde{\mathbf{x}}^* = \mathbf{x}^*$, so the $\mathbf{x}^{j,k}$ all converge to $\mathbf{x}^*|_{\Omega_j}$. \square

As we will show in Section 6, the supersolution constructed from Lemma 3.1 is generally very far from the solution to the global problem and should not be used as the initial guess in an actual computation. A much better choice would be to use the uniform initial guess $\mathbf{x}^0 = (0, 1/N, \dots, (N-1)/N, 1)^T$. We show that the alternating Schwarz method also converges to the exact solution with this initial guess.

For the steady-state problem, we would like to find a supersolution as well as a subsolution for the uniform initial guess. There, the residual function is simply h_i , so by putting $c = 2B/N^2$, Lemma 3.1 gives a \mathbf{u} that satisfies $h_i(\mathbf{u}) \geq h_i(\mathbf{x}^0)$, provided (3.1) holds, i.e.

$$\frac{\hat{m}}{N} + \frac{2B}{N} < \frac{\check{m}^2}{B} \implies N > \frac{B(\hat{m} + 2B)}{\check{m}^2}. \quad (4.6)$$

In other words, \mathbf{u} is a supersolution. Similarly, putting $c = -2B/N^2$ gives a subsolution, provided the same condition on N holds. We now arrive at our first main result.

THEOREM 4.2. *Let $N > \frac{B(\hat{m}+2B)}{\check{m}^2}$ with B , \hat{m} and \check{m} defined in (C1) and (C2). Suppose that for the uniform guess $\mathbf{x}^0 = (1/N, \dots, (N-1)/N)^T$, there are enough points in each subdomain Ω_j , $j = 1, \dots, N$, so that the alternating Schwarz method (4.3) applied to the nonlinear system H in (2.3) produces well-defined subdomain iterates $\mathbf{x}^{j,k}$ for all j and k . Then the subdomain iterates converge to $\mathbf{x}^*|_{\Omega_j}$, the restriction of the exact solution onto Ω_j .*

Proof. First, N is necessarily larger than the $B\hat{m}/\check{m}^2$ required for the existence of a solution by Theorem 3.2. In addition, the initial guess \mathbf{x}^0 belongs to the set D , since

$$x_i - x_{i-1} = \frac{1}{N} < \frac{\check{m}^2}{B(\hat{m} + 2B)} \leq \frac{\check{m}}{B}.$$

Moreover, the initial residual satisfies

$$|h_i(\mathbf{x}^0)| = \left| \tilde{M} \left(\frac{i-1}{N}, \frac{i}{N} \right) \cdot \frac{1}{N} - \tilde{M} \left(\frac{i}{N}, \frac{i+1}{N} \right) \cdot \frac{1}{N} \right| \leq \frac{2B}{N^2}.$$

As a result, by letting $c = \pm 2B/N^2$, we get

$$\frac{\hat{m}}{N} + N|c| = \frac{\hat{m} + 2B}{N} < \frac{\check{m}^2}{B},$$

so the hypotheses of Lemma 3.1 are satisfied. We thus conclude that there exist subsolutions and supersolutions $\tilde{\mathbf{x}}^0, \hat{\mathbf{x}}^0 \in D$ such that

$$H(\tilde{\mathbf{x}}^0) \leq H(\mathbf{x}^0) \leq H(\hat{\mathbf{x}}^0),$$

which implies

$$\tilde{\mathbf{x}}^0 \leq \mathbf{x}^0 \leq \hat{\mathbf{x}}^0. \quad (4.7)$$

We now prove inductively that for the alternating Schwarz (Gauss–Seidel) process, we have $\check{\mathbf{x}}^{j,k} \leq \mathbf{x}^{j,k} \leq \hat{\mathbf{x}}^{j,k}$ for $j = 1, \dots, J$ and for $k \geq 1$; this implies $\mathbf{x}^{j,k} \rightarrow \mathbf{x}^*|_{\Omega_j}$, since both $\check{\mathbf{x}}^{j,k}$ and $\hat{\mathbf{x}}^{j,k}$ converge to $\mathbf{x}^*|_{\Omega_j}$ by Theorem 4.1.

As an induction hypothesis, assume that there exists some $k \geq 0$ and $\ell \geq 1$ such that

$$\check{\mathbf{x}}^{j,k} \leq \mathbf{x}^{j,k} \leq \hat{\mathbf{x}}^{j,k} \quad 1 \leq j \leq J, \quad (4.8)$$

$$\check{\mathbf{x}}^{j,k+1} \leq \mathbf{x}^{j,k+1} \leq \hat{\mathbf{x}}^{j,k+1} \quad 1 \leq j \leq \ell - 1, \quad (4.9)$$

where for $\ell = 1$ the set in (4.9) is empty. Clearly, (4.8) and (4.9) hold for $k = 0$ and $\ell = 1$. We prove (4.9) for $\ell + 1$ as follows: by (4.3), we know that

$$\begin{aligned} H_\ell(\check{\mathbf{x}}^{\ell,k+1}; \check{x}_{a_\ell}^{\ell-1,k+1}, \check{x}_{b_\ell}^{\ell+1,k}) &= 0, \\ H_\ell(\mathbf{x}^{\ell,k+1}; x_{a_\ell}^{\ell-1,k+1}, x_{b_\ell}^{\ell+1,k}) &= 0, \\ H_\ell(\hat{\mathbf{x}}^{\ell,k+1}; \hat{x}_{a_\ell}^{\ell-1,k+1}, \hat{x}_{b_\ell}^{\ell+1,k}) &= 0, \end{aligned}$$

where, using the notation established in Section 4, H_ℓ refers to the ℓ -th subdomain problem for H . For convenience we write the boundary conditions as

$$\begin{aligned} \check{x}_{a_1}^{0,k+1} &= x_{a_1}^{0,k+1} = \hat{x}_{a_1}^{0,k+1} = 0, \\ \check{x}_{b_J}^{J+1,k} &= x_{b_J}^{J+1,k} = \hat{x}_{b_J}^{J+1,k} = 1. \end{aligned}$$

Now by (4.2), we have

$$0 = H_\ell(\mathbf{x}^{\ell,k+1}; x_{a_\ell}^{\ell-1,k+1}, x_{b_\ell}^{\ell+1,k}) = H_\ell(\hat{\mathbf{x}}^{\ell,k+1}; \hat{x}_{a_\ell}^{\ell-1,k+1}, \hat{x}_{b_\ell}^{\ell+1,k}) \leq H_\ell(\check{\mathbf{x}}^{\ell,k+1}; x_{a_\ell}^{\ell-1,k+1}, x_{b_\ell}^{\ell+1,k}).$$

Since H_ℓ is inverse isotone when the boundary conditions are fixed, this implies $\mathbf{x}^{\ell,k+1} \leq \hat{\mathbf{x}}^{\ell,k+1}$, and the same argument holds for the subsolution. Thus, (4.9) is verified for $\ell + 1$. This completes the proof of (4.9), and hence of (4.8), which in turn shows convergence of $\mathbf{x}^{j,k}$ to the solution $\mathbf{x}^*|_{\Omega_j}$. \square

REMARKS.

1. Suppose each subdomain has at least $B(\hat{m} + 2B)/\check{m}^2$ nodes, so that the first iteration of the alternating Schwarz method is defined. Then all subsequent iterates are guaranteed to be defined: since each subdomain solution is necessarily monotonic by the local equidistribution principle, see Lemma B.1, we have $0 \leq u_i \leq 1$ for all i . Thus, all interface values are between 0 and 1 at the next iteration, so the same number of mesh points suffices. In actual computations, however, one generally needs fewer grid points, see Section 6.
2. If a non-uniform initial guess is used, a result analogous to Theorem 4.2 also holds, except we need to replace $2B/N^2$ by $\max_i |h_i(\mathbf{x}^0)|$, and modify the number of grid points in (4.6) accordingly.
3. Remark 2 is important within the *MP* iteration procedure which alternately solves the physical and mesh DEs. The previous approximate mesh can be used as a good initial guess for the DD solution of next mesh DE solve. This generally reduces the number of DD iterations as the *MP* iteration proceeds. We provide an example of this in Section 6.

4.2. Convergence of the Parallel Schwarz iteration. Under the same hypothesis as in Theorem 4.1, we can show that the parallel Schwarz method also converges to the exact solution.

THEOREM 4.3. *Let N be large enough so that a unique solution $\mathbf{x}^* \in D \subset \mathbb{R}^{N+1}$ of the global problem exists. Let $\mathbf{x}^0 \in D \subset \mathbb{R}^{N+1}$ be a supersolution, i.e., $F(\mathbf{x}^0) \geq 0 = F(\mathbf{x}^*)$. Then the subdomain iterates $\mathbf{x}^{j,k}$ obtained from the classical parallel Schwarz method (4.4) with initial guess $\mathbf{x}^{j,0} = \mathbf{x}^0|_{\Omega_j}$ are well defined and satisfy*

$$\mathbf{x}^{j,k} \geq \mathbf{x}^{j,k+1} \geq \mathbf{x}^*|_{\Omega_j}$$

for all $k \geq 0$ and $\mathbf{x}^{j,k} \rightarrow \mathbf{x}^*|_{\Omega_j}$ as $k \rightarrow \infty$; that is, we have convergence from above. Similarly, if \mathbf{x}^0 is a subsolution, then $\mathbf{x}^{j,k} \uparrow \mathbf{x}^*|_{\Omega_j}$ as $k \rightarrow \infty$; that is, we have convergence from below.

Proof. The arguments for the supersolution case are similar to the proof of Theorem 4.1, but instead of defining the fractional iterates, we define \mathbf{x}^{k+1} by

$$x_i^{k+1} = \begin{cases} x_i^{j,k+1}, & \text{if } b_{j-1} \leq i \leq a_{j+1}, \\ \min\{x_i^{j-1,k+1}, x_i^{j,k+1}\}, & \text{if } a_j < i < b_{j-1}. \end{cases}$$

In other words, for points inside the overlap, we choose the smaller of the two solutions. We now show by induction that $F(\mathbf{x}^k) \geq 0$ and $\mathbf{x}^* \leq \mathbf{x}^{k+1} \leq \mathbf{x}^k$ for all $k \geq 0$. The base case for $F(\mathbf{x}^k) \geq 0$ is true for $k = 0$. Then we have

$$F_j(\mathbf{x}^k|_{\Omega_j}; x_{a_j}^k, x_{b_j}^k) \geq 0$$

for all j . For each index i , let Ω_j be the subdomain for which $x_i^{k+1} = x_i^{j,k+1}$. (The choice of j is unambiguous when i is outside the overlap, but when i is inside the overlap, we pick the j that yields the smaller value for $x_i^{j,k+1}$.) Then by construction, we have

$$f_i(\mathbf{x}^{j,k+1}) = 0 \leq f_i(\mathbf{x}^k).$$

By the inverse isotonicity of F_j , this implies $\mathbf{x}^{j,k+1} \leq \mathbf{x}^k|_{\Omega_j}$. Now $\mathbf{x}^{k+1}|_{\Omega_j} \leq \mathbf{x}^{j,k+1}$, since \mathbf{x}^{k+1} is defined as a pointwise minimum. By off-diagonal antitonicity of f_i , we see that

$$0 = f_i(\mathbf{x}^{j,k+1}) \leq f_i(\mathbf{x}^{k+1}).$$

Since this is true for all i , we conclude that $F(\mathbf{x}^{k+1}) \geq 0$, as well as

$$\mathbf{x}^* \leq \mathbf{x}^{k+1} \leq \mathbf{x}^k.$$

The argument for the subsolution is similar, except we should take the larger of the two subdomain values inside the overlap. \square

It is clear that for the steady-state problem, we can use the same super- and subsolution as in Theorem 4.2 to show that the parallel Schwarz method converges for the uniform guess $\mathbf{x}^0 = (1/N, \dots, (N-1)/N)^T$. We now state a similar result for the time-dependent problem, which is the second main result of our paper.

THEOREM 4.4. *Let $N > B(\hat{m} + \kappa)/\check{m}^2$, where B , \hat{m} , and \check{m} are defined in conditions (C1) and (C2) and κ is defined in (3.5). Suppose each subdomain Ω_j*

contains enough grid points so that the parallel and alternating Schwarz methods (4.4) and (4.3) applied to the nonlinear system G in (2.4) produce well-defined subdomain iterates. Then for the initial guess $\mathbf{x}^{[\nu]}$, both algorithms converge to the solution $\mathbf{x}^{[\nu+1]}$ of the time-dependent mesh generation problem (3.4).

Proof. The condition on N is the same as in Theorem 3.3, so a unique solution to the global problem exists. Consider once again the supersolution $\hat{\mathbf{u}}$ and subsolution $\check{\mathbf{u}}$ constructed in the proof of Theorem 3.3. As shown in the remark immediately after Theorem 3.3, $\hat{\mathbf{u}}$ and $\check{\mathbf{u}}$ satisfy

$$g_i(\check{\mathbf{u}}) \leq g_i(\mathbf{x}^{[\nu]}) \leq g_i(\hat{\mathbf{u}}).$$

Thus, by the M -function property of G , we have

$$\check{\mathbf{u}} \leq \mathbf{x}^{[\nu]} \leq \hat{\mathbf{u}}.$$

We can now mimic the proof of Theorem 4.2 to show that the parallel and alternating Schwarz methods (4.3) and (4.4) both converge to the unique solution \mathbf{x}^* for the initial guess $\mathbf{x}^{[\nu]}$. \square

REMARK. Suppose the $x_i^{[\nu]}$ are all between 0 and 1, and assume that $2(1-\theta)\gamma\hat{m} \leq 1$. (This is always true for $\theta = 1$, i.e., Backward Euler, for instance.) Then the discrete maximum principle (Lemma B.2) holds; as a result, if each subdomain has at least $B(\hat{m} + \kappa)/\hat{m}^2$ nodes, then every iteration of the alternating and parallel Schwarz methods is well defined, for the same reason as in the steady case.

5. Asymptotic convergence rate of the Schwarz iterations. In this section, we analyze the asymptotic convergence rate of the discrete parallel Schwarz method of Section 4 as $k \rightarrow \infty$, i.e., when \mathbf{x}^k is close to the solution \mathbf{x}^* . We only show the analysis for the stationary equidistribution problem

$$\phi_{i-1/2}(x_{i-1}, x_i) - \phi_{i+1/2}(x_i, x_{i+1/2}) = 0, \quad i = 1, \dots, N-1;$$

similar techniques can be used to analyze the time-dependent problem. Our first main result is the following:

THEOREM 5.1. *Suppose the conditions (C1)–(C3) hold. Then the parallel Schwarz method (4.4) for solving the stationary equidistribution problem (2.3) converges with the asymptotic convergence factor $\rho(Z)$, where $\rho(Z)$ denotes the spectral radius of the $(2J-2) \times (2J-2)$ matrix*

$$Z = \left[\begin{array}{cc|cc|} 0 & Q_1 & & & & & & \\ P_2 & 0 & 0 & R_2 & & & & \\ \hline S_2 & 0 & 0 & Q_2 & & & & \\ & & P_3 & 0 & 0 & R_3 & & \\ \hline & & S_3 & 0 & 0 & Q_3 & & \\ & & & & P_4 & 0 & 0 & R_4 \\ \hline & & & & \ddots & & \ddots & \\ \hline & & & & & & S_{J-1} & 0 & 0 & Q_{J-1} \\ & & & & & & & & P_J & 0 \end{array} \right]. \quad (5.1)$$

Moreover, the coefficients P_j , Q_j , R_j and S_j satisfy the bounds

$$\begin{aligned}\mu^{a_{j+1}-a_j} \cdot \frac{\mu^{-(b_j-a_{j+1})} - 1}{\mu^{-(b_j-a_j)} - 1} &\leq S_j \leq \mu^{-(a_{j+1}-a_j)} \cdot \frac{1 - \mu^{b_j-a_{j+1}}}{1 - \mu^{b_j-a_j}}, \\ \mu^{b_j-a_{j+1}} \cdot \frac{\mu^{-(a_{j+1}-a_j)} - 1}{\mu^{-(b_j-a_j)} - 1} &\leq Q_j \leq \mu^{-(b_j-a_{j+1})} \cdot \frac{1 - \mu^{a_{j+1}-a_j}}{1 - \mu^{b_j-a_j}}, \\ \mu^{b_{j-1}-a_j} \cdot \frac{\mu^{-(b_j-b_{j-1})} - 1}{\mu^{-(b_j-a_j)} - 1} &\leq P_j \leq \mu^{-(b_{j-1}-a_j)} \cdot \frac{1 - \mu^{b_j-b_{j-1}}}{1 - \mu^{b_j-a_j}}, \\ \mu^{b_j-b_{j-1}} \cdot \frac{\mu^{-(b_{j-1}-a_j)} - 1}{\mu^{-(b_j-a_j)} - 1} &\leq R_j \leq \mu^{-(b_j-b_{j-1})} \cdot \frac{1 - \mu^{b_{j-1}-a_j}}{1 - \mu^{b_j-a_j}},\end{aligned}$$

whenever $\mu \in (0, 1]$ is a constant such that for all $i = 1, \dots, N-1$, we have

$$\mu \left| \frac{\partial \phi_{i-1/2}}{\partial x_i} \right| \leq \left| \frac{\partial \phi_{i+1/2}}{\partial x_i} \right| \leq \frac{1}{\mu} \left| \frac{\partial \phi_{i-1/2}}{\partial x_i} \right|. \quad (5.2)$$

This result is analogous to the result in [22] for the continuous case: there, it is shown that for the *continuous* equidistribution problem, the error on the j th subdomain with respect to the error measure

$$E_j^k(\xi) = \int_{x_j(\xi)}^{x_j^k(\xi)} M(\tilde{x}) d\tilde{x}$$

satisfies

$$E_j^{k+1}(\xi) = \frac{1}{\xi_{b_j} - \xi_{a_j}} [(\xi - \xi_{a_j})E_{j+1}^k(\xi_{b_j}) + (\xi_{b_j} - \xi)E_{j-1}^k(\xi_{a_j})]. \quad (5.3)$$

If we now define the vector \mathbf{E}^k containing the errors at all interface points (ordered from left to right according to their location in the domain, see Fig. 4.1)

$$\mathbf{E}^k = (E_1^k(\xi_{a_2}), E_2^k(\xi_{b_1}), E_2^k(\xi_{a_3}), E_3^k(\xi_{b_2}), \dots, E_{J-1}^k(\xi_{a_J}), E_J^k(\xi_{b_{J-1}}))^T,$$

then we can use (5.3) to obtain $\mathbf{E}^{k+1} = \tilde{Z}\mathbf{E}^k$, where

$$\tilde{Z} = \begin{bmatrix} 0 & q_1 & & & & & & \\ p_2 & 0 & 0 & r_2 & & & & \\ s_2 & 0 & 0 & q_2 & & & & \\ & & p_3 & 0 & 0 & r_3 & & \\ & & s_3 & 0 & 0 & q_3 & & \\ & & & & p_4 & 0 & 0 & r_4 \\ & & & & & \ddots & \ddots & \ddots \\ & & & & & & s_{J-1} & 0 & 0 & q_{J-1} \\ & & & & & & & p_J & 0 & \end{bmatrix}, \quad (5.4)$$

with

$$p_j = \frac{\xi_{b_j} - \xi_{b_{j-1}}}{\xi_{b_j} - \xi_{a_j}}, \quad q_j = \frac{\xi_{a_{j+1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad r_j = \frac{\xi_{b_{j-1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad s_j = \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}}.$$

Our second main result of this section shows that as the grid is refined, the discrete iteration matrix converges to the continuous one above, subject to some smoothness assumptions on \tilde{M} :

THEOREM 5.2. *Let N be the number of degrees of freedom in the equidistribution problem. Suppose the conditions (C1)–(C3) hold, and assume in addition that the partial derivatives $\frac{\partial \tilde{M}}{\partial u}$ and $\frac{\partial \tilde{M}}{\partial v}$ are both Lipschitz continuous. Then the coefficients P_j , Q_j , R_j and S_j in Theorem 5.1 satisfy*

$$P_j \rightarrow \frac{\xi_{b_j} - \xi_{b_{j-1}}}{\xi_{b_j} - \xi_{a_j}}, \quad Q_j \rightarrow \frac{\xi_{a_{j+1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad R_j \rightarrow \frac{\xi_{b_{j-1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad S_j \rightarrow \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}},$$

as $N \rightarrow \infty$.

The next two subsections contain the proofs of these two theorems.

5.1. Proof of Theorem 5.1. First, let us write the parallel Schwarz method (4.4) in fixed point form in terms of the interface values. Suppose we arrange the interface values in the same order as in the continuous setting:

$$\mathbf{X}_\Gamma^k = (x_{a_2}^{1,k}, x_{b_1}^{2,k}, x_{a_3}^{2,k}, x_{b_2}^{3,k}, \dots, x_{a_J}^{J-1,k}, x_{b_{J-1}}^{J,k})^T.$$

Next, let Φ denote the fixed point mapping corresponding to the parallel Schwarz iteration, i.e., we have $\mathbf{X}_\Gamma^{k+1} = \Phi(\mathbf{X}_\Gamma^k)$. Then it is well known that the asymptotic convergence rate of the iteration is given by the spectral radius of the Jacobian $\frac{\partial \Phi}{\partial \mathbf{X}}$. Our first step is to determine the sparsity pattern of this Jacobian matrix.

LEMMA 5.3. *The Jacobian $\frac{\partial \Phi}{\partial \mathbf{X}}$ has the same sparsity pattern as the matrix Z in (5.1).*

Proof. If $\mathbf{Y} = \Phi(\mathbf{X})$ with $\mathbf{Y} = (y_{a_2}, y_{b_1}, \dots, y_{a_J}, y_{b_{J-1}})^T$, then for all subdomains $j = \{1, \dots, J\}$, the values y_{a_j} and y_{b_j} are defined implicitly by

$$\begin{aligned} y_{a_j} &= (\mathbf{y}^{j-1})_{a_j}, & \text{where } H_{j-1}(\mathbf{y}^{j-1}; x_{a_{j-1}}, x_{b_{j-1}}) &= 0, \\ y_{b_j} &= (\mathbf{y}^{j+1})_{b_j}, & \text{where } H_{j+1}(\mathbf{y}^{j+1}; x_{a_{j+1}}, x_{b_{j+1}}) &= 0, \end{aligned}$$

where the H_j are the subdomain objective functions, as defined in (4.1). Thus, y_{a_j} only depends on the values of $x_{a_{j-1}}$ and $x_{b_{j-1}}$, whereas y_{b_j} only depends on $x_{a_{j+1}}$, and $x_{b_{j+1}}$, so their derivatives with respect to the remaining variables are identically zero. This yields a Jacobian with the sparsity pattern shown in (5.1). \square

Our next step is to compute the individual coefficients of the Jacobian. Recall that the global problem is defined by $H(\mathbf{y}) = 0$, where the components h_i of H read

$$h_i(y_{i-1}, y_i, y_{i+1}) = \phi_{i-1/2}(y_{i-1}, y_i) - \phi_{i+1/2}(y_i, y_{i+1}), \quad a_{j-1} < i < b_{j-1}.$$

In order to match the results in the continuous case, we will consider a diagonally scaled version of the Jacobian, i.e., we let $Z = D \frac{\partial \Phi}{\partial \mathbf{X}} D^{-1}$, where

$$D = \text{diag} \left(\frac{\partial \phi_{a_2+1/2}}{\partial y_{a_2}}, \frac{\partial \phi_{b_1+1/2}}{\partial y_{b_1}}, \dots, \frac{\partial \phi_{a_J+1/2}}{\partial y_{a_J}}, \frac{\partial \phi_{b_{J-1}+1/2}}{\partial y_{b_{J-1}}} \right).$$

Comparing with the definition of Z in (5.1), we see that

$$\begin{aligned} P_j &= \left(\frac{\partial \phi_{a_j+1/2}}{\partial y_{a_j}} \right)^{-1} \frac{\partial y_{b_{j-1}}}{\partial x_{a_j}} \frac{\partial \phi_{b_{j-1}+1/2}}{\partial y_{b_{j-1}}}, & R_j &= \left(\frac{\partial \phi_{b_j+1/2}}{\partial y_{b_j}} \right)^{-1} \frac{\partial y_{b_{j-1}}}{\partial x_{b_j}} \frac{\partial \phi_{b_{j-1}+1/2}}{\partial y_{b_{j-1}}}, \\ S_j &= \left(\frac{\partial \phi_{a_j+1/2}}{\partial y_{a_j}} \right)^{-1} \frac{\partial y_{a_{j+1}}}{\partial x_{a_j}} \frac{\partial \phi_{a_{j+1}+1/2}}{\partial y_{a_{j+1}}}, & Q_j &= \left(\frac{\partial \phi_{b_j+1/2}}{\partial y_{b_j}} \right)^{-1} \frac{\partial y_{a_{j+1}}}{\partial x_{b_j}} \frac{\partial \phi_{a_{j+1}+1/2}}{\partial y_{a_{j+1}}}, \end{aligned} \tag{5.5}$$

where all derivatives are to be evaluated at the solution \mathbf{X}^* .

Using implicit differentiation, we obtain

$$\frac{\partial H_j}{\partial \mathbf{y}^j} \frac{\partial \mathbf{y}^j}{\partial x_{a_j}} + \frac{\partial H_j}{\partial x_{a_j}} = 0 \implies \frac{\partial \mathbf{y}^j}{\partial x_{a_j}} = - \left(\frac{\partial H_j}{\partial \mathbf{y}^j} \right)^{-1} \frac{\partial H_j}{\partial x_{a_j}}.$$

Thus, $\frac{\partial y_{a_j+1}}{\partial x_{a_j}}$ is simply the component of $\frac{\partial \mathbf{y}^j}{\partial x_{a_j}}$ corresponding to the location a_j+1 , and the other derivatives can be calculated similarly. The matrix $K_j := \frac{\partial H_j}{\partial \mathbf{y}^j}$ has a special structure: it is a tridiagonal matrix of the form

$$K_j = \begin{bmatrix} c_1 + d_1 & -c_2 & & & & \\ -d_1 & c_2 + d_2 & -c_3 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -d_{N_j-2} & c_{N_j-1} + d_{N_j-1} & -c_{N_j} & \\ & & & -d_{N_j-1} & c_{N_j} + d_{N_j} & \end{bmatrix}, \quad (5.6)$$

where $N_j = b_j - a_j - 1$ is the number of interior unknowns in subdomain j and

$$c_i = \frac{\partial \phi_{a_j+i-1/2}}{\partial y_{a_j+i}}, \quad d_i = -\frac{\partial \phi_{a_j+i+1/2}}{\partial y_{a_j+i}}.$$

Next, we notice that the vector $\frac{\partial H_j}{\partial x_{a_j}}$ has exactly one non-zero component at the first position. This is because the only component that depends on x_{a_j} corresponds to index $a_j + 1$, i.e., the first unknown in the nonlinear system. Its value can be calculated explicitly:

$$\frac{\partial h_{a_j+1}}{\partial x_{a_j}} = \frac{\partial \phi_{a_j+1/2}}{\partial x_{a_j}} = -d_0.$$

Thus, we see that the derivatives we want to compute are multiples of $(\mathbf{e}^q)^T K_j^{-1} \mathbf{e}^p$, where \mathbf{e}^q is the q -th standard basis vector (1 at position q and 0 elsewhere), q is the position of index a_j+1 in subdomain j , and $p \in \{1, N_j\}$ depending on whether we are taking the derivative with respect to x_{a_j} or x_{b_j} . In other words, we need to estimate individual entries of the inverse of a tridiagonal matrix. Estimates of this type exist in the literature, see e.g. [40] and [43]; unfortunately, those estimates are not precise enough for our purposes. In particular, to prove Theorem 5.2, we need the upper and lower bounds in 5.1 to converge to the same value as $\mu \rightarrow 1$. As a result, we will derive our own estimates with the help of the next lemma.

LEMMA 5.4. *Let K_j be defined in (5.6). Then for $q \in \{1, \dots, N_j\}$, we have*

$$(\mathbf{e}^q)^T K_j^{-1} \mathbf{e}^1 = \frac{\mu_1 \cdots \mu_{q-1}}{C_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right)}, \quad (\mathbf{e}^q)^T K_j^{-1} \mathbf{e}^{N_j} = \frac{\mu_{q+1}^{-1} \cdots \mu_{N_j}^{-1}}{D_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right)},$$

where $\mu_i = d_i/c_i$ and

$$C_q = 1 + \mu_{q-1} + \mu_{q-1}\mu_{q-2} + \cdots + \mu_{q-1} \cdots \mu_1, \\ D_q = 1 + \mu_{q+1}^{-1} + \mu_{q+1}^{-1}\mu_{q+2}^{-1} + \cdots + \mu_{q+1}^{-1} \cdots \mu_{N_j}^{-1}.$$

Proof. Let $\mathbf{w} = K_j^{-T} \mathbf{e}^q = (w_1, \dots, w_{N_j})^T$, so that for all p ,

$$(\mathbf{e}^q)^T K_j^{-1} \mathbf{e}^p = \mathbf{w}^T \mathbf{e}^p = w_p.$$

Then we have $K_j \mathbf{w} = \mathbf{e}^q$, which in component form reads

$$c_i(w_i - w_{i-1}) - d_i(w_{i+1} - w_i) = \begin{cases} 1, & i = q \\ 0, & \text{else,} \end{cases} \quad (5.7)$$

where for convenience we have defined $w_0 = w_{P+1} = 0$. This implies

$$w_i - w_{i-1} = \frac{d_i}{c_i}(w_{i+1} - w_i) \quad \text{for } i \neq q,$$

which allows us to unroll the recurrence to obtain

$$\begin{aligned} w_q &= w_q - w_0 = (w_q - w_{q-1}) + (w_{q-1} - w_{q-2}) + \dots + (w_1 - w_0) \\ &= (w_q - w_{q-1}) \left[1 + \frac{d_{q-1}}{c_{q-1}} + \frac{d_{q-1}d_{q-2}}{c_{q-1}c_{q-2}} + \dots + \frac{d_{q-1} \dots d_1}{c_{q-1} \dots c_1} \right] =: C_q(w_q - w_{q-1}). \end{aligned}$$

Unrolling the recurrence in the other direction gives another expression for w_q :

$$\begin{aligned} w_q &= w_q - w_{P+1} = (w_q - w_{q+1}) + (w_{q+1} - w_{q+2}) + \dots + (w_{N_j} - w_{N_j+1}) \\ &= (w_q - w_{q+1}) \left[1 + \frac{c_{q+1}}{d_{q+1}} + \frac{c_{q+1}c_{q+2}}{d_{q+1}d_{q+2}} + \dots + \frac{c_{q+1} \dots c_{N_j}}{d_{q+1} \dots d_{N_j}} \right] =: D_q(w_q - w_{q+1}). \end{aligned}$$

Now substituting the two expressions above into the q -th equation of (5.7) gives

$$\left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right) w_q = 1.$$

Thus, we get

$$\begin{aligned} w_{N_j} &= w_{N_j} - w_{N_j+1} = \frac{c_{q+1} \dots c_{N_j}}{d_{q+1} \dots d_{N_j}} (w_q - w_{q+1}) \\ &= \frac{c_{q+1} \dots c_{N_j}}{d_{q+1} \dots d_{N_j}} \frac{w_q}{D_q} = \frac{c_{q+1} \dots c_{N_j}}{d_{q+1} \dots d_{N_j}} \frac{1}{D_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right)}. \end{aligned}$$

Similarly, we have

$$w_1 = w_1 - w_0 = \frac{d_{q-1} \dots d_1}{c_{q-1} \dots c_1} \cdot \frac{1}{C_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right)}.$$

□

The above lemma can be generalized to obtain expressions for the other components of \mathbf{w} , but we will omit this since it is not required for our analysis.

EXAMPLE. For the case $c_i = d_i = 1$ for all i (i.e., the 1D Laplacian), we obtain $C_q = q$, $D_q = N_j - q + 1$, so

$$w_1 = \frac{N_j - q + 1}{N_j + 1}, \quad w_{N_j} = \frac{q}{N_j + 1}.$$

In particular, we have $w_1 + w_{N_j} = 1$.

We are now ready to prove Theorem 5.1.

Proof. (Theorem 5.1) For any invertible diagonal matrix D , the matrices $\frac{\partial \Phi}{\partial \mathbf{X}}$ and $D \frac{\partial \Phi}{\partial \mathbf{X}} D^{-1}$ have the same sparsity pattern, so by Lemma 5.3, it suffices to show that the coefficients P_j , Q_j , R_j and S_j satisfy the stated bounds. We prove the bounds in detail for S_j only, the others being similar.

Let $\mu \in (0, 1]$ be a constant satisfying (5.2), so that $\mu \leq d_i/c_i \leq \mu^{-1}$ for all i . Then for q fixed, the expressions in the statement of Lemma 5.4 imply that

$$\frac{1 - \mu^q}{1 - \mu} \leq C_q \leq \frac{\mu^{-q} - 1}{\mu^{-1} - 1}, \quad \frac{1 - \mu^{N_j - q + 1}}{1 - \mu} \leq D_q \leq \frac{\mu^{-(N_j - q + 1)} - 1}{\mu^{-1} - 1}.$$

The denominator of $w_1 = (\mathbf{e}^q)^T K_j^{-1} \mathbf{e}^1$ then satisfies

$$\begin{aligned} C_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right) &= c_q \left(1 + \frac{d_q}{c_q} \cdot \frac{C_q}{D_q} \right) \\ &\leq c_q \left(1 + \mu^{-1} \cdot \frac{\mu^{-q} - 1}{\mu^{-1} - 1} \cdot \frac{1 - \mu}{1 - \mu^{N_j - q + 1}} \right) = c_q \frac{\mu^{-(N_j + 1)} - 1}{\mu^{-(N_j - q + 1)} - 1}. \end{aligned}$$

On the other hand, we also have

$$C_q \left(\frac{c_q}{C_q} + \frac{d_q}{D_q} \right) \geq c_q \left(1 + \mu \cdot \frac{1 - \mu^q}{1 - \mu} \cdot \frac{\mu^{-1} - 1}{\mu^{-(N_j - q + 1)} - 1} \right) = c_q \frac{1 - \mu^{N_j + 1}}{1 - \mu^{N_j - q + 1}}.$$

This implies

$$\frac{d_1 \cdots d_{q-1}}{c_1 \cdots c_q} \cdot \frac{\mu^{-(N_j - q + 1)} - 1}{\mu^{-(N_j + 1)} - 1} \leq w_1 \leq \frac{d_1 \cdots d_{q-1}}{c_1 \cdots c_q} \cdot \frac{1 - \mu^{N_j - q + 1}}{1 - \mu^{N_j + 1}}.$$

Now let q be the local index of a_{j+1} within the j th subdomain, i.e., $q = a_{j+1} - a_j$. Then by definition, we have

$$\frac{\partial y_{a_j+1}}{\partial x_{a_j}} = d_0 w_1.$$

But $\frac{\partial \phi_{a_j+1/2}}{\partial y_{a_j}} = -d_0$ and $\frac{\partial \phi_{a_{j+1}+1/2}}{\partial y_{a_{j+1}}} = -d_q$, so in fact

$$S_j = \frac{d_q}{d_0} \cdot d_0 w_1 = d_q w_1,$$

from which it follows that

$$\frac{d_1 \cdots d_q}{c_1 \cdots c_q} \cdot \frac{\mu^{-(N_j - q + 1)} - 1}{\mu^{-(N_j + 1)} - 1} \leq S_j \leq \frac{d_1 \cdots d_q}{c_1 \cdots c_q} \cdot \frac{1 - \mu^{N_j - q + 1}}{1 - \mu^{N_j + 1}}.$$

The bound on the ratios d_i/c_i then yields

$$\mu^q \cdot \frac{\mu^{-(N_j - q + 1)} - 1}{\mu^{-(N_j + 1)} - 1} \leq S_j \leq \mu^{-q} \cdot \frac{1 - \mu^{N_j - q + 1}}{1 - \mu^{N_j + 1}}.$$

Substituting $N_j = b_j - a_j - 1$ and $q = a_{j+1} - a_j$ yields the required bounds on S_j . The bounds on P_j , Q_j and R_j are obtained similarly. \square

5.2. Proof of Theorem 5.2. The bounds in Theorem 5.1 depend on the value of μ that can be used. To understand the behavior of the algorithm as the number of grid points increases, we need the following lemma.

LEMMA 5.5. *Let K_j be defined in (5.6) and N be the number of degrees of freedom in the equidistribution problem. Suppose the conditions (C1)–(C3) hold, and assume in addition that the partial derivatives $\frac{\partial \tilde{M}}{\partial u}$ and $\frac{\partial \tilde{M}}{\partial v}$ are both Lipschitz continuous. Then for N large enough, there exists a constant $\omega > 0$ independent of N such that if we let*

$$\mu = 1 - \frac{\omega}{N^2},$$

then we have $\mu \leq d_i/c_i \leq \mu^{-1}$ for all i .

Proof. Let $(y_i)_{i=0}^N$ be the solution of the equidistribution problem. From the condition (C3), we see that $\sum_{i=1}^N \tilde{M}(y_{i-1}, y_i)(y_i - y_{i-1})$ is simply a Riemann sum for the integral $\int_0^1 M(y) dy$, so it converges to the integral as $N \rightarrow \infty$. Thus, the Riemann sum is bounded for all N by a constant, say $\omega_1 > 0$. However, from the relation

$$\tilde{M}(y_0, y_1)(y_1 - y_0) = \tilde{M}(y_1, y_2)(y_2 - y_1) = \cdots = \tilde{M}(y_{N-1}, y_N)(y_N - y_{N-1}),$$

we see that for all i , we have

$$\tilde{M}(y_{i-1}, y_i)(y_i - y_{i-1}) = \frac{1}{N} \sum_{i=1}^N \tilde{M}(y_{i-1}, y_i)(y_i - y_{i-1}) \leq \frac{\omega_1}{N}.$$

This implies

$$0 < y_i - y_{i-1} < \frac{\omega_1}{\tilde{m}N}, \quad i = 1, \dots, N.$$

Next, we have by definition

$$d_i = -\frac{\partial \phi_{i+1/2}}{\partial y_i} = \frac{\partial}{\partial y_i} \left(\tilde{M}(y_i, y_{i+1})(y_i - y_{i+1}) \right) = \tilde{M}(y_i, y_{i+1}) + \frac{\partial \tilde{M}}{\partial u} \Bigg|_{(y_i, y_{i+1})} (y_i - y_{i+1})$$

$$c_i = \frac{\partial \phi_{i-1/2}}{\partial y_i} = \frac{\partial}{\partial y_i} \left(\tilde{M}(y_{i-1}, y_i)(y_i - y_{i-1}) \right) = \tilde{M}(y_{i-1}, y_i) + \frac{\partial \tilde{M}}{\partial v} \Bigg|_{(y_{i-1}, y_i)} (y_i - y_{i-1})$$

which means

$$c_i \geq \tilde{m} - B \frac{\omega_1}{\tilde{m}N} \geq \frac{\tilde{m}}{2} > 0$$

for N large enough, and the exact same bound holds for d_i . Moreover, we have

$$d_i - c_i = \underbrace{\tilde{M}(y_i, y_{i+1}) - \tilde{M}(y_{i-1}, y_i)}_{(*)} + \frac{\partial \tilde{M}}{\partial u} \Bigg|_{(y_i, y_{i+1})} (y_i - y_{i+1}) - \frac{\partial \tilde{M}}{\partial v} \Bigg|_{(y_{i-1}, y_i)} (y_i - y_{i-1}).$$

The term in (*) can be written as

$$\begin{aligned} \tilde{M}(y_i, y_{i+1}) - \tilde{M}(y_{i-1}, y_i) &= \tilde{M}(y_i, y_{i+1}) - \tilde{M}(y_i, y_i) + \tilde{M}(y_i, y_i) - \tilde{M}(y_{i-1}, y_i) \\ &= \frac{\partial \tilde{M}}{\partial v} \Bigg|_{(y_i, \zeta^+)} (y_{i+1} - y_i) + \frac{\partial \tilde{M}}{\partial u} \Bigg|_{(\zeta^-, y_i)} (y_i - y_{i-1}), \end{aligned}$$

where $y_{i-1} < \zeta^- < y_i$ and $y_i < \zeta^+ < y_{i+1}$. This implies

$$\begin{aligned} |d_i - c_i| &\leq \left| \left(\frac{\partial \tilde{M}}{\partial v} \Big|_{(y_i, \zeta^+)} (y_{i+1} - y_i) - \frac{\partial \tilde{M}}{\partial v} \Big|_{(y_{i-1}, y_i)} (y_i - y_{i-1}) \right) \right| \\ &\quad + \left| \left(\frac{\partial \tilde{M}}{\partial u} \Big|_{(\zeta^-, y_i)} (y_i - y_{i-1}) - \frac{\partial \tilde{M}}{\partial u} \Big|_{(y_i, y_{i+1})} (y_{i+1} - y_i) \right) \right| \\ &\leq \frac{\omega_1}{N} \left| \left(\frac{1}{\tilde{M}(y_i, y_{i+1})} \frac{\partial \tilde{M}}{\partial v} \Big|_{(y_i, \zeta^+)} - \frac{1}{\tilde{M}(y_{i-1}, y_i)} \frac{\partial \tilde{M}}{\partial v} \Big|_{(y_{i-1}, y_i)} \right) \right| \\ &\quad + \frac{\omega_1}{N} \left| \left(\frac{1}{\tilde{M}(y_{i-1}, y_i)} \frac{\partial \tilde{M}}{\partial u} \Big|_{(\zeta^-, y_i)} - \frac{1}{\tilde{M}(y_i, y_{i+1})} \frac{\partial \tilde{M}}{\partial u} \Big|_{(y_i, y_{i+1})} \right) \right| \end{aligned}$$

Thus, within each absolute value term, we have a difference of bounded Lipschitz functions evaluated at points that are at most $\omega_1/(\tilde{m}N)$ apart, which means there exists a constant ω_2 such that

$$|d_i - c_i| \leq \omega_2/N^2.$$

This implies

$$\frac{d_i}{c_i} = 1 + \frac{d_i - c_i}{c_i} \geq 1 - \frac{\omega_2}{N^2} \cdot \frac{2}{\tilde{m}}$$

for N large enough. Hence, putting $\omega = 2\omega_2/\tilde{m}$ and $\mu = 1 - \omega/N^2$ leads to $d_i/c_i \geq \mu$. Similarly, we have

$$\frac{c_i}{d_i} = 1 + \frac{c_i - d_i}{d_i} \geq 1 - \frac{\omega_2}{N^2} \cdot \frac{2}{\tilde{m}} = \mu,$$

so we have $c_i/d_i \geq \mu$, or equivalently, $d_i/c_i \leq \mu^{-1}$, which completes the proof. \square

We are finally ready to prove the second main result of this section.

Proof. (Theorem 5.2) We need to show that as $N \rightarrow \infty$, we have

$$P_j \rightarrow \frac{\xi_{b_j} - \xi_{b_{j-1}}}{\xi_{b_j} - \xi_{a_j}}, \quad Q_j \rightarrow \frac{\xi_{a_{j+1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad R_j \rightarrow \frac{\xi_{b_{j-1}} - \xi_{a_j}}{\xi_{b_j} - \xi_{a_j}}, \quad S_j \rightarrow \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}}.$$

We only prove the result for S_j , the other coefficients being similar. From Theorem 5.1, we have

$$\mu^{a_{j+1}-a_j} \cdot \frac{\mu^{-(b_j-a_{j+1})} - 1}{\mu^{-(b_j-a_j)} - 1} \leq S_j \leq \mu^{-(a_{j+1}-a_j)} \cdot \frac{1 - \mu^{b_j-a_{j+1}}}{1 - \mu^{b_j-a_j}}, \quad (5.8)$$

so if we can show that both the upper and lower bounds tend to $(\xi_{b_j} - \xi_{a_{j+1}})/(\xi_{b_j} - \xi_{a_j})$ as $N \rightarrow \infty$, then we are done. Indeed, from Lemma 5.5, we can choose $\mu = 1 - \omega/N^2$, so that

$$\mu^N = \left(1 - \frac{\omega}{N^2}\right)^N \rightarrow 1 \quad \text{as } N \rightarrow \infty. \quad (5.9)$$

Thus, by letting $\lambda = \mu^N$ and noting that $a_j = N\xi_{a_j}$, $b_j = N\xi_{b_j}$, we can rewrite (5.8) as

$$\lambda^{\xi_{a_{j+1}} - \xi_{a_j}} \cdot \frac{\lambda^{-(\xi_{b_j} - \xi_{a_{j+1}})} - 1}{\lambda^{-(\xi_{b_j} - \xi_{a_j})} - 1} \leq S_j \leq \lambda^{-(\xi_{a_{j+1}} - \xi_{a_j})} \cdot \frac{1 - \lambda^{\xi_{b_j} - \xi_{a_{j+1}}}}{1 - \lambda^{\xi_{b_j} - \xi_{a_j}}},$$

where the exponents are now independent of N . Hence, by taking the limit as $\lambda \rightarrow 1$, we obtain

$$\lambda^{(\xi_{a_{j+1}} - \xi_{a_j})} \rightarrow 1, \quad \frac{\lambda^{-(\xi_{b_j} - \xi_{a_{j+1}})} - 1}{\lambda^{-(\xi_{b_j} - \xi_{a_j})} - 1} \rightarrow \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}}, \quad \frac{1 - \lambda^{\xi_{b_j} - \xi_{a_{j+1}}}}{1 - \lambda^{\xi_{b_j} - \xi_{a_j}}} \rightarrow \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}}.$$

Thus, we have

$$S_j \rightarrow \frac{\xi_{b_j} - \xi_{a_{j+1}}}{\xi_{b_j} - \xi_{a_j}},$$

as required. \square

REMARK. We now see the importance of the N^2 in the estimate of Lemma 5.5: if we only had the estimate $\mu = 1 - \omega/N$, then (5.9) would give $\mu^N \rightarrow e^{-\omega}$, and the upper and lower bounds of S_j would not converge to the same limit.

6. Numerical Results.

6.1. A first example. We begin by considering the parallel Schwarz algorithm (4.4) applied to the steady mesh generation problem with

$$M(x) = \sqrt{1 + u_x^2} \quad \text{and} \quad u(x) = \frac{1 - e^{Rx}}{1 - e^R}.$$

For $R = 3$ the solution u exhibits a boundary layer at $x = 1$ and the mesh should have increased density in the location of the layer. For x outside of $[0, 1]$ we have extended M by a constant.

The single domain solution, computed by solving the BVP with a Newton iteration is shown in Figure 6.1. We see a concentration of mesh points near $x = 1$. The supersolution and subsolution are computed using N points where N is calculated by equation (3.1) using the bounds $\hat{m} = 3.4$, $\check{m} = 1$ and $B = 9.1$. From equation (4.6) this gives a requirement of $N = 198$ points. The shooting approach outlined in the proof of Lemma 3.1 is used along with a simple bisection approach to find s^* . We plot the supersolution, subsolution, the uniform initial guess and the single domain solution in Figure 6.2.

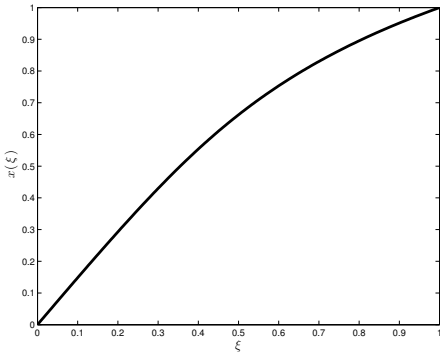


FIG. 6.1. Single domain solution $x(\xi)$ obtained with a Newton iteration for the boundary layer problem

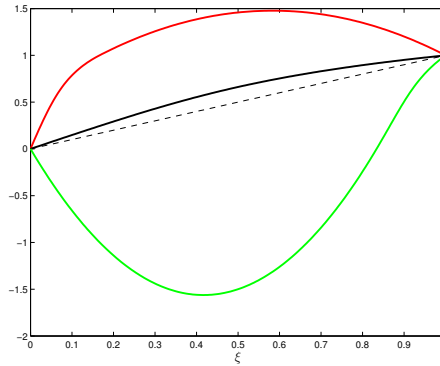


FIG. 6.2. Supersolution and subsolution for the boundary layer example.

In Figures 6.3 and 6.4 we illustrate the iterates obtained using the parallel DD iteration on two subdomains starting from the supersolution and subsolution. The plot shows iterates converging to the single domain solution.

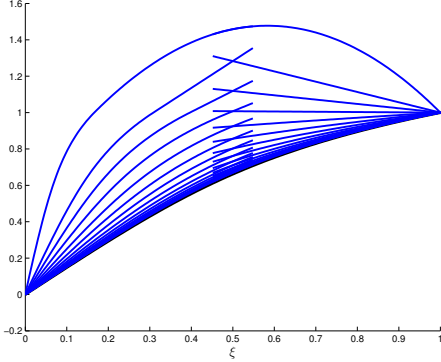


FIG. 6.3. DD iterations from the supersolution.

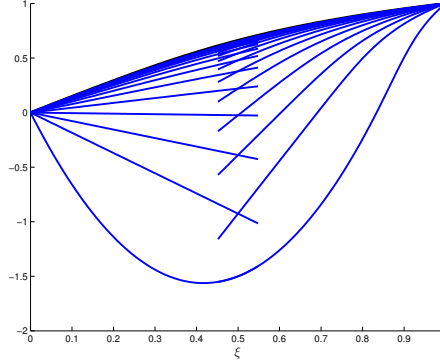


FIG. 6.4. DD iterations from the subsolution.

In Figures 6.5 and 6.6 we illustrate the convergence histories on subdomains one and two with varying amounts of overlap K . We see the convergence improves as the amount of overlap increases. This is consistent with the asymptotic convergence rate obtained in Section 5.

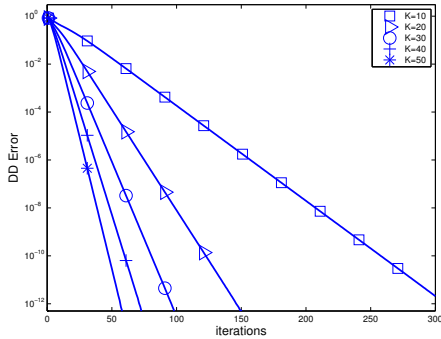


FIG. 6.5. The $\|\cdot\|_\infty$ difference between the DD iterates on subdomain one.

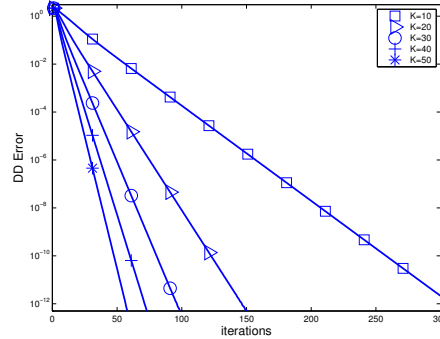


FIG. 6.6. The $\|\cdot\|_\infty$ difference between the DD iterates on subdomain two.

We see very close agreement between the theoretical and numerically estimated convergence rates in Table 6.1. The numerically estimated convergence rates are obtained by averaging the ratio of errors over five iterations as we near convergence.

	Overlap				
	10	20	30	40	50
ρ	0.8302	0.6791	0.5526	0.4482	0.3625
$\hat{\rho}$	0.8329	0.6791	0.5526	0.4482	0.3625

TABLE 6.1

Expected ($\hat{\rho}$) and obtained (ρ) convergence rates as a function of overlap for the boundary layer problem.

6.2. A more challenging problem. We now consider a more difficult problem by choosing

$$M(x) = 1 + 40e^{-(x-0.25)^2/0.01} + 10e^{-(x-0.75)^2/0.05}.$$

This choice of M will concentrate mesh nodes where M is large, i.e. near $x = 0.25$ and $x = 0.75$, as shown in Figure 6.7. A quick calculation gives the parameters $\tilde{m} = 1.7$, $\hat{m} = 6$ and $B = 45$. Our theory indicates that using $N = 1496$ points is sufficient to guarantee the existence of a subsolution and supersolution for this problem.

The single domain Newton solver requires 6 iterations and the resulting solution is shown in Figure 6.7.

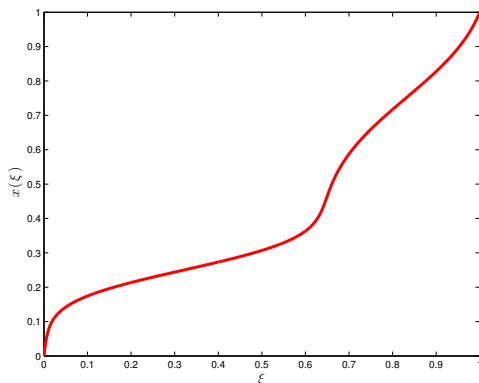


FIG. 6.7. *Single domain solution computed using a Newton iteration.*

In Figure 6.8 we illustrate the supersolution and subsolution for this problem along with the single domain solution (in solid black). The uniform mesh is shown in a dashed black line. The DD iterates obtained starting from the supersolution are shown in Figure 6.9.

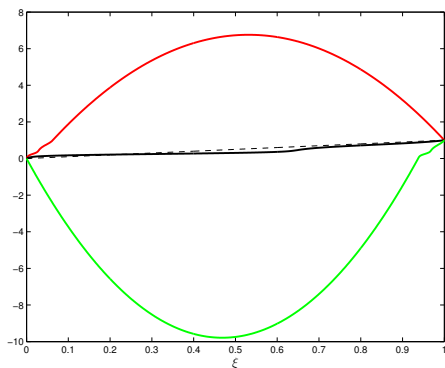


FIG. 6.8. *Supersolution and subsolution for problem two along with the single domain solution.*

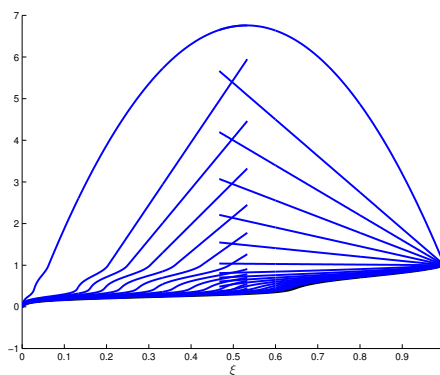


FIG. 6.9. *DD iterates from supersolution for problem two.*

The theory presented in Section 4 suggests that the sequence of meshes, $x^{j,k}$ obtained from an initial uniform mesh, should lie between $\tilde{x}^{j,k}$ and $\hat{x}^{j,k}$, the sequence of

meshes obtained by starting the DD iteration with the subsolution \tilde{x} and supersolution \hat{x} . This is confirmed in Table 6.2 where we report the differences between $x^{1,k}$ and $\hat{x}^{1,k}$ and $\tilde{x}^{1,k}$ at the interface ξ_{a_2} . Indeed these differences remain positive and are monotonically decreasing to zero as k increases.

k	1	2	3	4	5	6	7	8	9	10
$(\hat{x}^{1,k} - x^{1,k})(\xi_{a_2})$	6.18	5.31	4.58	3.86	3.30	2.74	2.32	1.89	1.56	1.23
$(x^{1,k} - \tilde{x}^{1,k})(\xi_{a_2})$	10.26	8.76	6.68	5.59	3.99	3.16	1.93	1.39	0.36	0.18

TABLE 6.2

The sequence of meshes starting with a uniform initial guess are trapped between the converging sequence of meshes starting from the subsolution and supersolution.

In Table 6.3 we see if we can recover the asymptotic rate of convergence as N increases. As we increase N we adjust the number of mesh points in the overlap region so as keep the overlap width approximately constant. For each value of N we estimate the numerical rate of convergence by averaging the ratio of errors over five iterations as we near convergence. We also record the theoretical convergence rate. The theoretical convergence rate $\hat{\rho}$ changes as the overlap width varies slightly with increasing N .

	N					
	25	50	75	100	125	150
ρ	0.5233	0.4789	0.4663	0.4605	0.4569	0.4547
$\hat{\rho}$	0.5102	0.4756	0.4648	0.4596	0.4564	0.4545

TABLE 6.3

Expected ($\hat{\rho}$) and obtained (ρ) convergence rates as a function of overlap for the boundary layer problem.

6.3. A note on robustness. As M becomes more difficult the solution of the nonlinear BVP for the mesh transformation using a Newton iteration becomes quite difficult. Indeed a very good initial guess is often required. Here we show that the nonlinear Schwarz DD method, with a sufficient number of subdomains, gives additional robustness, and is able to recover the single domain solution from a uniform initial guess in a situation where the single domain Newton iteration fails.

Here we consider a more challenging M given by

$$M(x) = 1 + 30e^{-(x-0.25)^2/0.01} + 40e^{-(x-0.75)^2/0.05}. \quad (6.1)$$

In Table 6.4 we show the difficulty that the single domain Newton solve has for this nonlinear BVP (starting from a uniform initial guess) for various values of N near 100. During the DD solution we alter the global number of mesh points slightly to ensure that each interior subdomain has the same number of degrees of freedom. We see that for N near 100 the Newton iteration has difficulty converging on a single domain. Indeed for $N = 101$ and $N = 105$ no convergence is obtained even after 100 Newton iterations. For the other values of N the convergence only seems to occur by happenchance!

N	100	101	103	104	105
# iters	17	NC	67	26	NC

TABLE 6.4

The number of single domain Newton iterations versus the number of mesh points N for the mesh generation BVP with M given in (6.1).

We now consider dividing the unit interval into subdomains and using a parallel Schwarz iteration to obtain the global solution. Within each subdomain, we use Newton's method to solve the nonlinear subdomain problems. In Table 6.5 we give the number of Newton iterations required for each subdomain solve during the first DD iteration as a function of the number of subdomains used. We can see that as the number of subdomains increases the nonlinear problems solved on the subdomains become easier. For example, using $N = 101$ points the global Newton iteration fails. Using 4 subdomains convergence is achieved, albeit with great difficulty, requiring 43 Newton iterations on subdomain 1. With 8 subdomains, and again 101 global mesh points, only 4-7 Newton iterations are required on each subdomain. The DD iteration itself converges for both 4 and 8 subdomains and hence we are able to recover the global solution which is not possible with the global Newton iteration.

S	N	Ω_1	Ω_2	Ω_3	Ω_4	Ω_5	Ω_6	Ω_7	Ω_8
1	100	17							
2	101	75	5						
3	101	31	6	5					
4	101	43	6	5	5				
5	100	43	6	5	4	5			
6	101	8	5	5	5	4	5		
7	103	7	5	6	5	4	4	5	
8	101	7	5	5	4	5	4	4	5

TABLE 6.5

The number of Newton iterations required on each subdomain as a function of the number of subdomains, S , using approximately $N = 100$ mesh points.

In Tables 6.6 and 6.7 we repeat the experiment above but now with approximately 150 global mesh points. A similar pattern emerges. The single domain Newton iteration fails or requires many iterations. With a sufficient number of subdomains the global solution is recovered.

N	150	151	152	153	155	157
# iters	65	28	NC	NC	NC	76

TABLE 6.6
iters versus N

S	N	Ω_1	Ω_2	Ω_3	Ω_4	Ω_5	Ω_6	Ω_7	Ω_8
1	150	65							
2	151	NC	5						
3	152	21	6	5					
4	153	32	6	5	5				
5	150	9	6	5	4	5			
6	155	8	4	5	5	4	5		
7	152	7	5	5	5	4	4	5	
8	157	7	5	5	4	5	4	4	4

TABLE 6.7

The number of Newton iterations required on each subdomain as a function of the number of subdomains, S , using approximately $N = 150$ mesh points.

6.4. DD within the MP framework. We conclude our presentation of numerical results with a short illustration of how the domain decomposition method performs within the *MP* alternating procedure described in the introduction.

Consider the solution of the two-point boundary value problem

$$-\epsilon u'' + u' = 1, u(0) = u(1) = 0,$$

with $\epsilon = 0.05$. The physical solution

$$u(x) = x - \frac{1 - e^{x/\epsilon}}{1 - e^{1/\epsilon}},$$

has a sharp boundary layer near $x = 1$.

The *MP* procedure proceeds as follows. Pick an (often uniform) initial mesh $X^{(0)}$. For $k = 0, 1, \dots$

- Solve the physical PDE for $U^{(k)}$ on the mesh $X^{(k)}$.
- Solve for $X^{(k+1)}$ using the mesh DE, where the monitor function $M(x)$ is calculated based on $X^{(k)}$ and $U^{(k)}$, see below.

In the mesh solve step, we solve

$$M((x_i + x_{i-1})/2)(x_i - x_{i-1}) - M((x_i + x_{i+1})/2)(x_{i+1} - x_i) = 0 \quad (6.2)$$

for the new mesh $\{x_i\}$, where the continuous function M is constructed by interpolating the ordered pairs $(\hat{x}_{i-1/2}, M_{i-1/2})$ where the $M_{i-1/2}$ are computed using the previous mesh $\{\hat{x}\}$ and physical solution on that mesh. We choose

$$M_{i-1/2} = \sqrt{1 + \left(\frac{u_i - u_{i-1}}{\hat{x}_i - \hat{x}_{i-1}} \right)^2}$$

as an approximation to M at the half nodes $\hat{x}_{i-1/2} = (\hat{x}_i + \hat{x}_{i-1})/2$.

The convergence of the *MP* iteration procedure, starting with an uniform initial mesh with $N = 51$ mesh points, is shown in Figure 6.10.

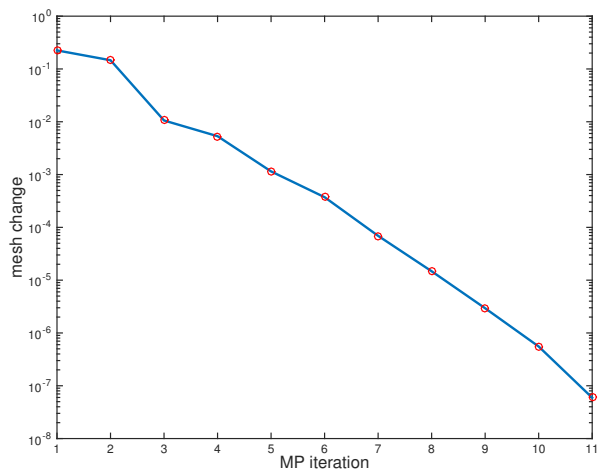


FIG. 6.10. *MP iteration convergence for a boundary layer problem.*

The initial uniform mesh and the solution to the boundary value problem is shown in Figure 6.11.

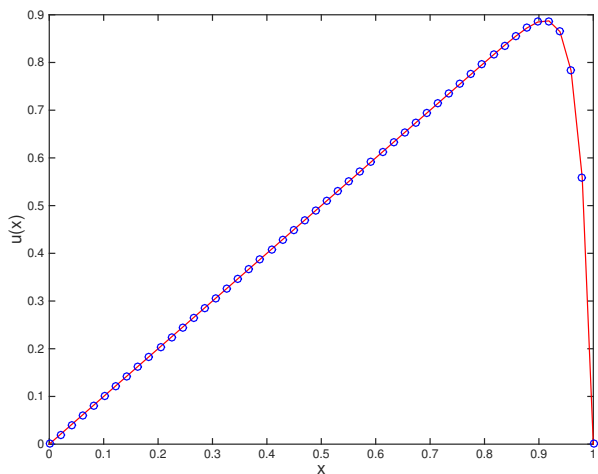


FIG. 6.11. *Solution of the BVP and an initial uniform mesh.*

The solution of the BVP and the first mesh $X^{(1)}$ obtained by solving (6.2) during the *MP* process described above is shown in Figure 6.12.

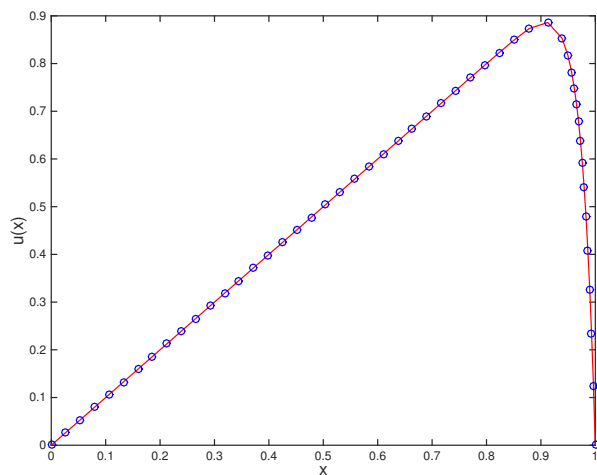


FIG. 6.12. Solution of the BVP and the first mesh obtained during our MP procedure.

During the *MP* procedure one could linearize the mesh solves by simply evaluating the mesh density function at the previously computed mesh and solution. In other words, instead of constructing \bar{M} by interpolation and evaluating at $(x_i + x_{i-1})/2$ and $(x_i + x_{i+1})/2$, one could instead use the values of $M_{i-1/2}$ and $M_{i+1/2}$ directly. This results in a different mesh $X^{(1)}$, shown in Figure 6.13. Comparing Figures 6.12 and 6.13, we see that the initial mesh quality using the linearized variant is clearly not as good as the *MP* procedure we describe.

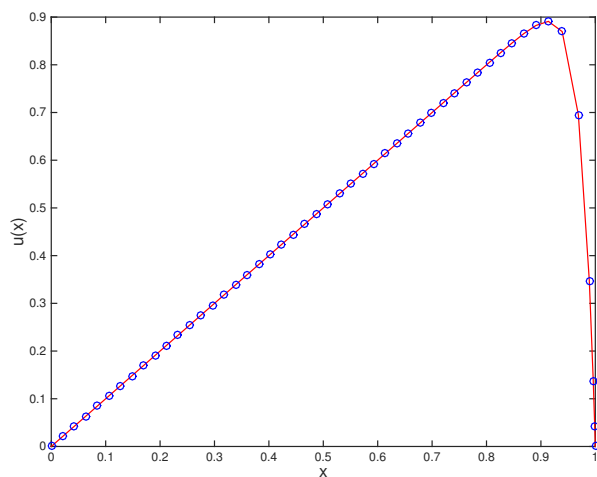


FIG. 6.13. Solution of the BVP and the first mesh obtained during the linearized MP procedure.

During each *MP* iteration, the previous mesh $\{\hat{x}_i\}$ is used as initial guess for the DD iteration to obtain the new mesh. Using two subdomains, the number of DD iterations required for the difference in DD approximations to agree within 10^{-5} for the first 7 *MP* iterations are given in Table 6.8.

<i>MP</i> iteration	1	2	3	4	5	6	7
# <i>DD</i> iterations	63	49	54	33	19	22	11

TABLE 6.8

The number of *DD* iterations for each *MP* iteration.

We can see that starting the *DD* iteration for (6.2) using the mesh found during the previous *MP* iteration generally reduces the number of *DD* iterations required to converge to the specified tolerance.

7. Conclusions. In this paper, we have shown that when the integral form of the moving mesh PDE is discretized using a quadrature rule, the resulting nonlinear algebraic system may not have a solution if the number of grid points is too small. Using *M*-function theory, we have proved the existence and uniqueness of the solution of the discrete problem when the number of grid points is large enough. Moreover, we have shown that the parallel and alternating Schwarz methods converge to the unique solution, so they can be used to solve the discrete solution in parallel. In fact, as the mesh is refined, these methods behave asymptotically the same way as their continuous counterparts, with the same iteration matrix and the same convergence rates. This is also confirmed by our numerical results. Our numerical experiments show that the domain decomposition algorithms are often more robust than the methods that try to find the single-domain solution all at once, such as Newton’s method. And finally, we show that the *DD* approach can effectively be used within the *MP* iteration to solve the coupled problem.

We believe that this *M*-function theory can be applied to analyze the behavior of domain decomposition methods when applied to other nonlinear PDEs; this is a subject of ongoing investigation. Ongoing work also focuses on understanding the behavior when optimized conditions are used, see [22], as well as for the 2D mesh equation in [33].

Appendix A. M-function characterization. To prove Lemma 2.8, we will first need the following geometrical fact.

PROPOSITION A.1. *Let \mathbb{D} be an open convex subset of \mathbb{R}^n . For two given points $x, y \in \mathbb{D}$, let $r > 0$ be a constant such that the open balls $B_r(x)$ and $B_r(y)$, of radius r and centered at x and y respectively, are contained in \mathbb{D} . Then for any intermediate point $p = (1 - t)x + ty$ with $t \in [0, 1]$, we have $B_r(p) \subset \mathbb{D}$.*

Proof. Let $v \in \mathbb{R}^n$ be a vector with $\|v\| < r$, so that $x + v \in \mathbb{D}$ and $y + v \in \mathbb{D}$. Then by the convexity of \mathbb{D} , we have

$$(1 - s)(x + v) + s(y + v) = (1 - s)x + sy + v \in \mathbb{D} \quad \forall s \in [0, 1].$$

Choosing $s = t$ shows that $p + v \in \mathbb{D}$. Since v is arbitrary, we conclude that $B_r(p) \subset \mathbb{D}$. \square

Now we proceed to prove Lemma 2.8 which we restate here.

LEMMA 2.8 *Let \mathbb{D} be a convex and open subset of \mathbb{R}^n . Assume $F : \mathbb{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is off-diagonally antitone, and that for any $\mathbf{x} \in \mathbb{D}$, the functions $q_i : S_i \subset \mathbb{R} \rightarrow \mathbb{R}^n$ defined as*

$$q_i(\tau) = \sum_{j=1}^n f_j(\mathbf{x} + \tau \mathbf{e}^i), \quad i = 1, \dots, n, \quad \text{with} \quad S_i = \{\tau : \mathbf{x} + \tau \mathbf{e}^i \in \mathbb{D}\},$$

are isotone. Finally, assume for every i that either q_i is strictly isotone, or there exists a node $j = j(i)$ such that a path $j \rightsquigarrow i$ exists and q_j is strictly isotone. Then F is an M -function.

Proof. To prove Lemma 2.8 we need to show that F is inverse isotone, so let $\mathbf{x} \neq \mathbf{y} \in \mathbb{D}$ be such that $F(\mathbf{x}) \leq F(\mathbf{y})$. We must now show that $\mathbf{x} \leq \mathbf{y}$. Let $r > 0$ be such that $B_r(x)$ and $B_r(y)$ are both contained in \mathbb{D} . Moreover, let $\mathbf{x}_s = (1-s)\mathbf{y} + s\mathbf{x}$, so that $s \in [0, 1] \implies \mathbf{x}_s \in \mathbb{D}$. Now define $t \in (0, 1]$ so that

$$\frac{1}{t} = \max \left\{ 1, \left\lceil \frac{N \max_i |x_i - y_i|}{r} \right\rceil \right\} =: M \in \mathbb{N}.$$

Then the points $\mathbf{y} = \mathbf{x}_0, \mathbf{x}_t, \mathbf{x}_{2t}, \dots, \mathbf{x}_{Mt} = \mathbf{x}$ all satisfy $B_r(\mathbf{x}_{kt}) \subset \mathbb{D}$ for $k = 0, \dots, M$ by Proposition A.1. Now define the sets

$$N^- = \{i \in \mathbb{N} : y_i < x_i\}, \quad N^+ = \{i \in \mathbb{N} : y_i \geq x_i\}.$$

Let us assume $N^- \neq \emptyset$, and let i_1, \dots, i_m be nodes of N^- and set

$$\mathbf{z}^j = (x_{i_j} - y_{i_j})\mathbf{e}^{i_j}, \quad j = 1, \dots, m, \quad \mathbf{z} = \sum_{j=1}^m \mathbf{z}^j,$$

where \mathbf{e}^{i_j} is the i_j -th coordinate vector. Noting that $\|\mathbf{z}^j\| \leq r/N$, we see that the sequence

$$\mathbf{y} + t\mathbf{z}^1, \mathbf{y} + t\mathbf{z}^1 + t\mathbf{z}^2, \dots, \mathbf{y} + t \sum_{j=1}^m \mathbf{z}^j = \mathbf{y} + t\mathbf{z}$$

all belong to \mathbb{D} , since \mathbb{D} is convex. Then by the isotonicity of the q_i , we have

$$\sum_{j=1}^n f_j(\mathbf{y}) \leq \sum_{j=1}^n f_j(\mathbf{y} + t\mathbf{z}^1) \leq \sum_{j=1}^n f_j(\mathbf{y} + t\mathbf{z}^1 + t\mathbf{z}^2) \leq \dots \leq \sum_{j=1}^n f_j(\mathbf{y} + t\mathbf{z}). \quad (\text{A.1})$$

Splitting the last sum into sums over N^- and N^+ , we use the antitonicity of the f_j to obtain

$$\sum_{j=1}^n f_j(\mathbf{y} + t\mathbf{z}) = \sum_{j \in N^-} f_j(\mathbf{y} + t\mathbf{z}) + \sum_{j \in N^+} f_j(\mathbf{y}) \leq \sum_{j \in N^-} f_j(\mathbf{x}_t) + \sum_{j \in N^+} f_j(\mathbf{y}). \quad (\text{A.2})$$

Thus, we have

$$\sum_{j \in N^-} f_j(\mathbf{y}) \leq \sum_{j \in N^-} f_j(\mathbf{x}_t). \quad (\text{A.3})$$

We can now continue this argument to obtain

$$\sum_{j \in N^-} f_j(\mathbf{y}) \leq \sum_{j \in N^-} f_j(\mathbf{x}_t) \leq \sum_{j \in N^-} f_j(\mathbf{x}_{2t}) \leq \dots \leq \sum_{j \in N^-} f_j(\mathbf{x}_{Mt}) = \sum_{j \in N^-} f_j(\mathbf{x}). \quad (\text{A.4})$$

But since we had assumed $F(\mathbf{x}) \leq F(\mathbf{y})$, we must have $f_j(\mathbf{y}) = f_j(\mathbf{x})$ for all $j \in N^-$, so that all the inequalities in (A.4) (and hence (A.1), (A.2) and (A.3)) are in fact

equalities. This, in turn, shows that q_i cannot be strictly isotone for any $i \in N^-$. Thus, there must exist a strict link (i, i') such that $i \in N^+$, $i' \in N^-$, i.e., we have

$$f_i(\mathbf{y}) > f_i(\mathbf{y} + t(x_{i'} - y_{i'})\mathbf{e}^{i'})$$

as well as

$$f_j(\mathbf{y}) \geq f_j(\mathbf{y} + t(x_{i'} - y_{i'})\mathbf{e}^{i'}) \quad \text{whenever } j \in N^+.$$

Thus,

$$\begin{aligned} \sum_{j=1}^n f_j(\mathbf{y}) &= \sum_{j \in N^-} f_j(\mathbf{y} + t\mathbf{z}) + \sum_{j \in N^+} f_j(\mathbf{y}) \\ &> \sum_{j \in N^-} f_j(\mathbf{y} + t\mathbf{z}) + \sum_{j \in N^+} f_j(\mathbf{y} + t(x_{i'} - y_{i'})\mathbf{e}^{i'}) \\ &\geq \sum_{j \in N^-} f_j(\mathbf{y} + t\mathbf{z}) + \sum_{j \in N^+} f_j(\mathbf{y} + t\mathbf{z}) \\ &= \sum_{j=1}^n f_j(\mathbf{y} + t\mathbf{z}) = \sum_{j=1}^n f_j(\mathbf{y}), \end{aligned}$$

which is a contradiction. Thus N^- must be empty, so that F is inverse isotone, and hence an M -function. \square

Appendix B. Discrete maximum principles. In this section, we prove two discrete maximum principles, which are used to deduce the number of grid points necessary to guarantee the existence of subdomain iterates.

LEMMA B.1 (Steady mesh generation). *Let $\{u_i\}_{i=0}^N$ be such that $u_0 = u_L$, $u_N = u_R$ and*

$$\tilde{M}(u_i, u_{i+1})(u_{i+1} - u_i) - \tilde{M}(u_{i-1}, u_i)(u_i - u_{i-1})$$

for $1 \leq i \leq N - 1$. Then

$$\min(u_L, u_R) \leq u_i \leq \max(u_L, u_R).$$

Proof. We have

$$\tilde{M}(u_0, u_1)(u_1 - u_0) = \tilde{M}(u_1, u_2)(u_2 - u_1) = \cdots = \tilde{M}(u_{N-1}, u_N)(u_N - u_{N-1}).$$

Since $M(u, v) \geq \tilde{m} > 0$ for all u and v , the differences $u_i - u_{i-1}$ all have the same sign, so there can be no internal minima or maxima. Thus the maximum and minimum values are $u_0 = u_L$ and $u_N = u_R$ respectively. \square

LEMMA B.2 (Time-dependent mesh generation). *Let $\{u_i\}_{i=0}^N$ be such that $u_0 = u_L$, $u_N = u_R$ and*

$$\begin{aligned} &u_i - \theta\gamma(\tilde{M}(u_i, u_{i+1})(u_{i+1} - u_i) - \tilde{M}(u_{i-1}, u_i)(u_i - u_{i-1})) \\ &= x_i^{[\nu]} + (1 - \theta)\gamma(\tilde{M}(x_i^{[\nu]}, x_{i+1}^{[\nu]})(x_{i+1}^{[\nu]} - x_i^{[\nu]}) - \tilde{M}(x_{i-1}^{[\nu]}, x_i^{[\nu]})(x_i^{[\nu]} - x_{i-1}^{[\nu]})). \end{aligned} \quad (\text{B.1})$$

for $1 \leq i \leq N - 1$. If $2(1 - \theta)\gamma\hat{m} \leq 1$, then

$$\min(u_L, u_R, \min_i x_i) \leq u_i \leq \max(u_L, u_R, \max_i x_i).$$

Proof. We show only the maximum, the minimum being similar. Let

$$\begin{aligned} m_1 &= \tilde{M}(u_{i-1}, u_i), & m_2 &= \tilde{M}(u_i, u_{i+1}), \\ m_3 &= \tilde{M}(x_{i-1}^{[\nu]}, x_i^{[\nu]}), & m_4 &= \tilde{M}(x_i^{[\nu]}, x_{i+1}^{[\nu]}), \end{aligned}$$

so that $\tilde{m} \leq m_j \leq \hat{m}$ for $j = 1, 2, 3, 4$, and let $\bar{x} = \max_i x_i^{[\nu]}$. Then for $i = 1, \dots, N - 1$, the right-hand side of (B.1) can be bounded by

$$\begin{aligned} & x_i^{[\nu]} + (1 - \theta)\gamma(m_4(x_{i+1}^{[\nu]} - x_i^{[\nu]}) - m_3(x_i^{[\nu]} - x_{i-1}^{[\nu]})) \\ &= \underbrace{(1 - \gamma(1 - \theta)(m_3 + m_4))}_{\geq 0} x_i^{[\nu]} + \gamma(1 - \theta)m_3 x_{i-1}^{[\nu]} + \gamma(1 - \theta)m_4 x_{i+1}^{[\nu]} \\ &\leq (1 - \gamma(1 - \theta)(m_3 + m_4))\bar{x} + \gamma(1 - \theta)m_3\bar{x} + \gamma(1 - \theta)m_4\bar{x} = \bar{x}. \end{aligned}$$

On the other hand, suppose u_i is the maximum value, i.e., $u_i \geq u_j$ for all j . Then we either have $i = 0$ or $i = N$, or the left-hand side of (B.1) satisfies

$$u_i - \theta\gamma \underbrace{(m_1(u_{i-1} - u_i) + m_2(u_{i+1} - u_i))}_{\leq 0} \geq u_i.$$

Thus, if the maximum occurs at $1 \leq i \leq N - 1$, then we must have $u_i \leq \bar{x}$. \square

REFERENCES

- [1] Christoph Augustin and Olaf Steinbach. FETI methods for the simulation of biological tissues. In *Domain Decomposition Methods in Science and Engineering XX*, pages 503–510. Springer, 2013.
- [2] D. Bennequin, M. J. Gander, and L. Halpern. A homographic best approximation problem with application to optimized Schwarz waveform relaxation. *Math. of Comp.*, 78(265):185–232, 2009.
- [3] Heiko Berninger and Oliver Sander. Substructuring of a Signorini-type problem and Robin’s method for the Richards equation in heterogeneous soil. *Computing and visualization in science*, 13(5):187–205, 2010.
- [4] Alexander Bihlo and Ronald D. Haynes. Parallel stochastic methods for PDE based grid generation. *Comput. & Math. Appl.*, 68:804–820, Oct 2014.
- [5] Morton Bjorhus. *On Domain Decomposition, Subdomain Iteration, and Waveform Relaxation*. PhD thesis, Norwegian Institute of Technology, The University of Trondheim, 1995.
- [6] Igor P. Boglaev. Iterative algorithms of domain decomposition for the solution of a quasilinear elliptic problem. *J. Comput. Appl. Math.*, 80(2):299–316, 1997.
- [7] Hermann G. Burchard. Splines (with optimal knots) are better. *Applicable Anal.*, 3:309–319, 1973/74.
- [8] Xiao-Chuan Cai. Additive Schwarz algorithms for parabolic convection-diffusion equations. *Numer. Math.*, 60(1):41–61, 1991.
- [9] Xiao-Chuan Cai. Multiplicative Schwarz methods for parabolic problems. *SIAM J. Sci. Comput.*, 15(3):587–603, 1994. Iterative methods in numerical linear algebra (Copper Mountain Resort, CO, 1992).
- [10] Xiao-Chuan Cai and Maksymilian Dryja. Domain decomposition methods for monotone nonlinear elliptic problems. In *Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993)*, volume 180 of *Contemp. Math.*, pages 21–27. Amer. Math. Soc., Providence, RI, 1994.

- [11] Xiao-Chuan Cai and David E. Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SIAM J. Sci. Comput.*, 24(1):183–200 (electronic), 2002.
- [12] Xiao-Chuan Cai, David E. Keyes, and Leszek Marcinkowski. Non-linear additive Schwarz preconditioners and application in computational fluid dynamics. *Internat. J. Numer. Methods Fluids*, 40(12):1463–1470, 2002. LMS Workshop on Domain Decomposition Methods in Fluid Mechanics (London, 2001).
- [13] L. Collatz. Aufgaben monotoner Art. *Arch. Math.*, 3:366–376, 1952.
- [14] Thu-Huyen Dao, Michael Ndjinga, and Frédéric Magoulès. A Schur complement method for compressible Navier-Stokes equations. In *Domain Decomposition Methods in Science and Engineering XX*, pages 543–550. Springer, 2013.
- [15] C. de Boor. Good approximation by splines with variable knots. II. In *Conference on the Numerical Solution of Differential Equations (Univ. Dundee, Dundee, 1973)*, pages 12–20. Lecture Notes in Math., Vol. 363. Springer, Berlin, 1974.
- [16] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, 37(2):296–311, 1997.
- [17] Martin J. Gander. A waveform relaxation algorithm with overlapping splitting for reaction diffusion equations. *Numerical Linear Algebra with Applications*, 6:125–145, 1998.
- [18] Martin J. Gander. Optimized Schwarz methods. *SIAM J. Numer. Anal.*, 44(2):699–731, 2006.
- [19] Martin J. Gander and Laurence Halpern. Absorbing boundary conditions for the wave equation and parallel computing. *Math. Comp.*, 74(249):153–176, 2004.
- [20] Martin J. Gander and Laurence Halpern. Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. *SIAM J. Numer. Anal.*, 45(2):666–697, 2007.
- [21] Martin J. Gander, Laurence Halpern, and Frédéric Nataf. Optimal Schwarz waveform relaxation for the one dimensional wave equation. *SIAM J. Numer. Anal.*, 41(5):1643–1681, 2003.
- [22] Martin J. Gander and Ronald D. Haynes. Domain decomposition approaches for mesh generation via the equidistribution principle. *SIAM J. Numer. Anal.*, 50(4):2111–2135, 2012.
- [23] Martin J. Gander and Christian Rohde. Overlapping Schwarz waveform relaxation for convection dominated nonlinear conservation laws. *SIAM J. Sci. Comp.*, 27(2):415–439, 2005.
- [24] Martin J. Gander and Andrew M. Stuart. Space time continuous analysis of waveform relaxation for the heat equation. *SIAM J.*, 19:2014–2031, 1998.
- [25] Eldar Giladi and Herbert B. Keller. Space time domain decomposition for parabolic problems. *Numerische Mathematik*, 93(2):279–313, 2002.
- [26] Florian Häberlein and Laurence Halpern. Optimized Schwarz waveform relaxation for nonlinear systems of parabolic type. In *Domain Decomposition Methods in Science and Engineering XXI*. Springer, 2014.
- [27] Florian Häberlein, Laurence Halpern, and Anthony Michel. Newton-Schwarz optimised waveform relaxation Krylov accelerators for nonlinear reactive transport. In *Domain Decomposition Methods in Science and Engineering XX*, pages 387–394. Springer, 2013.
- [28] Ronald D. Haynes and Alexander J. Howse. Alternating Schwarz methods for partial differential equation-based mesh generation. *Int. J. Comput. Math.*, 92, 2015.
- [29] Weizhang Huang. Practical aspects of formulation and solution of moving mesh partial differential equations. *J. Comput. Phys.*, 171(2):753–775, 2001.
- [30] Weizhang Huang, Yuhe Ren, and Robert D. Russell. Moving mesh partial differential equations (MMPDES) based on the equidistribution principle. *SIAM J. Numer. Anal.*, 31(3):709–730, 1994.
- [31] Weizhang Huang and Robert D. Russell. A moving collocation method for solving time dependent partial differential equations. *Appl. Numer. Math.*, 20(1-2):101–116, 1996. Workshop on the method of lines for time-dependent problems (Lexington, KY, 1995).
- [32] Weizhang Huang and Robert D. Russell. *Adaptive Moving Mesh Methods*, volume 174 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2011.
- [33] Weizhang Huang and David M. Sloan. A simple adaptive grid method in two dimensions. *SIAM J. Sci. Comput.*, 15(4):776–797, 1994.
- [34] Feng-Nan Hwang and Xiao-Chuan Cai. A parallel nonlinear additive Schwarz preconditioned inexact Newton algorithm for incompressible Navier-Stokes equations. *J. Comput. Phys.*, 204(2):666–691, 2005.
- [35] Feng-Nan Hwang and Xiao-Chuan Cai. A class of parallel two-level nonlinear Schwarz preconditioned inexact Newton algorithms. *Comput. Methods Appl. Mech. Engrg.*, 196(8):1603–1611, 2007.
- [36] P.-L. Lions. On the Schwarz alternating method. I. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987)*, pages 1–42. SIAM, Philadelphia, PA, 1988.
- [37] S. H. Lui. On linear monotone iteration and Schwarz methods for nonlinear elliptic PDEs.

- Numer. Math.*, 93(1):109–129, 2002.
- [38] Shiu-Hong Lui. On monotone and Schwarz alternating methods for nonlinear elliptic PDEs. *M2AN Math. Model. Numer. Anal.*, 35(1):1–15, 2001.
 - [39] Tarek P. A. Mathew. *Domain decomposition methods for the numerical solution of partial differential equations*, volume 61 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2008.
 - [40] Gérard Meurant. A review on the inverse of symmetric tridiagonal and block tridiagonal matrices. *SIAM Journal on Matrix Analysis and Applications*, 13(3):707–728, 1992.
 - [41] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.
 - [42] Serguei Ovtchinnikov and Xiao-Chuan Cai. One-level Newton-Krylov-Schwarz algorithm for unsteady non-linear radiation diffusion problem. *Numer. Linear Algebra Appl.*, 11(10):867–881, 2004.
 - [43] R. Peluso and T. Politi. Some improvements for two-sided bounds on the inverse of diagonally dominant tridiagonal matrices. *Linear Algebra and its Applications*, 330(1):1–14, 2001.
 - [44] Werner C. Rheinboldt. On M -functions and their application to nonlinear Gauss-Seidel iterations and to network flows. *J. Math. Anal. Appl.*, 32:274–307, 1970.
 - [45] Oliver Sander. Coupling geometrically exact Cosserat rods and linear elastic continua. In *Domain Decomposition Methods in Science and Engineering XX*, pages 443–450. Springer, 2013.
 - [46] Xue-Cheng Tai and Magne Espedal. Rate of convergence of some space decomposition methods for linear and nonlinear problems. *SIAM J. Numer. Anal.*, 35(4):1558–1570, 1998.
 - [47] Andrea Toselli and Olof Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2005.
 - [48] X. Xu, W. Huang, and R. D. Russell. Convergence of de Boor’s algorithm for the generation of equidistributing meshes. *IMA Journal of Numerical Analysis*, 31(2):580–596, March 2011.