- -

Domain Decomposition Approaches for Mesh Generation via the Equidistribution Principle

- -

Gander, Martin Jakob; Haynes, Ronald D.

DOMAIN DECOMPOSITION APPROACHES FOR MESH GENERATION VIA THE EQUIDISTRIBUTION PRINCIPLE

MARTIN J. GANDER AND RONALD D. HAYNES

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. We propose in this paper several Schwarz domain decomposition algorithms for this task. We then study in detail the convergence properties of these algorithms applied to the nonlinear mesh PDE in one spatial dimension. We prove convergence for classical transmission conditions, and optimal and optimized variants for the generation of steady equidistributing grids. A classical, parallel, Schwarz algorithm is presented and analysed for the generation of time dependent (moving) equidistributing grids. We conclude our study with numerical experiments.

Key words. domain decomposition, Schwarz methods, moving meshes, equidistribution

AMS subject classifications. 65M55, 65M15, 65Y05, 65M50, 65L50, 65N50

1. Introduction. The computation of solutions of partial differential equations (PDEs) which vary over disparate space and time scales often benefit from the use of non–uniform meshes chosen to adapt to the local solution structures. There are three broad categories of adaptive methods for PDEs. Static refinement techniques, such as h–refinement, locally coarsen or refine the mesh according to some *a posteriori* error estimate. In the finite element setting, the order of the basis polynomials is varied resulting in p–refinement methods. These techniques have been combined to give hp adaptive finite element methods, see [13] for a recent review. In this article, we consider a class of r–refinement or moving mesh methods. As the solution evolves this approach adapts an initial (typically uniform) grid by *relocating* the mesh nodes while keeping the number of mesh nodes and mesh topology fixed. The mesh is determined by solving a so–called *moving mesh PDE (MMPDE)* which is coupled to the physical PDE of interest.

A standard way to perform mesh adaptation in space is to use the equidistribution principle (EP), see [12]. Given some positive measure $M(t,x,u)$ of the error or difficulty in the solution $u(t,x)$, where $x \in \Omega_p$, the physical domain, and $t \in [0,T]$, the EP in one spatial dimension requires that the mesh points, $x_i, i = 0, \dots, N$, satisfy

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

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

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

where $\theta(t) \equiv \int_0^1 M(t,\tilde{x},u)\,d\tilde{x}$ is the total error in the solution. The function M is referred to as the monitor or mesh density function; typically it is chosen so that M is large where we expect the error in the computed solution to be large. Enforcing (EP) concentrates mesh points in regions where the error is large. It follows directly,

by differentiating the continuous form of (EP), that the required mesh transformation satisfies, for all t, the nonlinear differential equation

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

As discussed in [33], it is possible to derive a linear grid generation formulation. However, in general, the nonlinear description is more robust — less prone to mesh singularities for example. For time dependent problems, (1.1) coupled with a physical parabolic PDE results in a challenging numerical problem. Semi–discretizing in space, or applying the (moving) method of lines, results in an index–2 DAE system which is stiff and ill–conditioned [1]. As a result, (1.1) is often relaxed to require equidistribution at time $t + \tau$, resulting in a number of parabolic MMPDEs [32], for example, the mesh may be found by solving

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

The relaxation parameter τ is chosen in practice (see [30]) so that the mesh evolves at a rate commensurate with that of the solution $u(t, x)$.

In practice, the MMPDE is coupled through the monitor function to the physical PDE, $u_t = \mathcal{L}(u)$, for example, where $\mathcal{L}$ is a spatial differential operator. Moving mesh methods have been developed and analyzed by numerous groups; see for example [14, 45] for CFD problems, [48, 47] for flow and magnetohydrodynamics and [38, 39] for discontinuous Galerkin formulations and 1D stability and convergence analysis. Thorough recent reviews of grid generation by the equidistribution principle and moving mesh methods for partial differential equations can be found in [8, 33].

Numerically, there have been numerous suggestions as to how to solve the coupled system of PDEs

$$\frac{\partial u}{\partial t} = \mathcal{L}(u), \quad \frac{\partial x}{\partial t} = \frac{1}{\tau} \frac{\partial}{\partial \xi} \left(M(t, x(\xi, t), u) \frac{\partial x}{\partial \xi} \right), \tag{1.2}$$

for $x \in \Omega_p$, $t \in (0, T]$. In [31], Huang and Russell describe the solver MOVCOL in which (1.2) is semidiscretized in space using a collocation approach and the system of ordinary differential equations is solved in time using DASSL [43]. The mesh x^k and solution u^k at time t_k are integrated forward to the new time level t_{k+1} simultaneously by solving a single nonlinear system for x^{k+1} and u^{k+1}. In two or more space dimensions this approach may be cost prohibitive. Indeed Babuska and Rheinboldt [2] note that it may not be necessary to solve for the mesh and solution to the same level of accuracy. Motivated by this idea, Mackenzie et al. [4] suggest an iterative approach. To move forward in time from t_k to t_{k+1}, the mesh and physical solution are solved alternately until the meshes agree to some tolerance.

Here, our ultimate goal is the parallel solution of PDEs on (moving) equidistributing grids, ie. the parallel solution of (1.2). We propose introducing spatial parallelism by solving (1.2) through a domain decomposition (DD) approach. The DD approach attempts to solve a PDE by a divide and conquer philosophy – partitioning the spatial domain Ω into overlapping or non–overlapping subdomains and thus reformulating the original PDE into an equivalent coupled system of PDEs. The coupling is provided by transmission conditions which are designed to appropriately match solutions on neighboring subdomains. The matching is imposed in such a way as to maintain

as much coarse grain granularity in the computation as possible, lending itself to implementation on distributed memory architecture using MPI (for example).

The literature involving DD approaches for elliptic problems is vast, see the books [46, 40]. 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, applying the alternating or parallel Schwarz method to the sequence of elliptic problems which result upon semi–discretization (in time) of the PDE [9, 10], or the more recent approach, decomposing the space–time domain and applying the Schwarz waveform relaxation algorithm [6, 24, 25]. For both approaches, optimized variants have been developed, see [18] and the references therein for the first, and [21, 19, 20, 5] for the second approach. Even if the physical PDE is linear, the coupled system (1.2) will be highly nonlinear. Comparably, much less has been written about DD applied to nonlinear PDEs, see [11, 37, 36, 35, 15, 7, 44] for the steady case, and [17, 23] for evolution problems.

There are several natural ways to couple DD and r–refinement strategies to design an effective solver. Indeed this work follows the experimental papers [29, 28, 27]. The attempts thus far can be roughly differentiated by the choice of spatial variable in which to apply DD: the physical coordinate $x \in \Omega_p$ or the computational coordinate $\xi \in \Omega_c$. In [29], the overlapping classical Schwarz waveform relaxation algorithm, using Dirichlet transmission conditions, is applied to the coupled system of mesh and physical PDEs (1.2) in the physical coordinates. The partitioning of the physical coordinate is fixed but mesh points are given freedom to move within the subdomain and the overlap region. The fixed subdomain boundary prevents mesh points from moving from the subdomain into the overlap region or vice–versa. The more recent papers [28, 27] were primarily interested in the mesh generation problem only. That is, the solution u was assumed to be known. Again working in the physical coordinate system, [28] fixed the boundary of the extended subdomain (subdomain plus overlap), allowing mesh points to move in and out of the overlap region. More recently, [27], introduced an overlapping decomposition of the computational coordinate ξ, ie. sub-domains were specified by a fixed number of mesh points and not a particular region of physical space. In a time dependent scenario this gives rise to moving subdomains. It is this later approach, DD in the computational coordinate, we will adopt here.

In this paper we focus on the theoretical analysis of solving for grids via the equidistribution principle by various parallel DD strategies. Suppose $u(t, x)$ is a specified function on $\Omega_p \times (0, T)$ with $\Omega_p := (0, 1)$. From (MMPDE5) a moving (equidistributing) grid may be found by solving the nonlinear diffusion equation on $\Omega_c := (0, 1)$,

$$\frac{\partial x}{\partial t} = \frac{\partial}{\partial \xi}\left(M(x, t)\frac{\partial x}{\partial \xi} \right), \quad x(\xi, 0) = x_0(\xi), \quad x(0, t) = 0, \quad x(1, t) = 1. \qquad (1.3)$$

If u is a time independent function on $\Omega_p = (0, 1)$, then from (1.1) an equidistributing mesh transformation is found by solving the nonlinear two–point boundary value problem (BVP)

$$\frac{d}{d\xi}\left(M(x)\frac{dx}{d\xi} \right) = 0, \quad x(0) = 0, \quad x(1) = 1, \qquad (1.4)$$

for the mesh transformation $x(\xi) : \Omega_c \to \Omega_p$. The analysis of various DD approaches to solve (1.3) and (1.4) will be the focus of this paper.

3

In the next section we state and prove results concerning the well–posedness of the subdomain problems defined in Sections 3, 4, 5 and 7. Section 3 proposes nonlinear and linearized classical Schwarz iterations for (1.4) and corresponding convergence results. In Section 4 an optimal Schwarz algorithm, providing convergence in a finite number of iterations, is presented. An optimized Schwarz variant, which approximates the optimal Schwarz algorithm, is discussed in Section 5. The extension of the classical Schwarz method to an arbitrary (finite) number of subdomains is given in Section 6. Section 7 provides a first look at the Schwarz algorithms for mesh generation in the time–dependent case. Numerical results are given to support our analysis in Section 8 and we conclude in Section 9 with a summary and plans for future work.

2. Mathematical Preliminaries. The DD methods presented in Sections 3, 4, 5 and 7 require the solution of (1.4) and (1.3) subject to Dirichlet and Robin type boundary conditions. We now collect various results which establish the well–posedness and implicit representations of the solutions of these subdomain problems.

We begin by considering (1.4) on an arbitrary subdomain $\xi \in (a, b) \subset \Omega_c = (0, 1)$ subject to Dirichlet boundary conditions

$$\frac{d}{d\xi}\left(M(x)\frac{dx}{d\xi}\right) = 0, \quad x(a) = \gamma_a, \quad x(b) = \gamma_b. \tag{2.1}$$

This will be the subdomain problem of interest for the nonlinear Schwarz algorithms of Sections 3 and 6.

LEMMA 2.1. *If M is differentiable and bounded away from zero and infinity, ie. there exists $\check{m}$ and $\hat{m}$ such that $0 < \check{m} \le M(x) \le \hat{m} < \infty$ for all x, then the BVP (2.1) has a unique solution given implicitly by*

$$\int_{\gamma_a}^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = \frac{\xi - a}{b - a} \int_{\gamma_a}^{\gamma_b} M(\tilde{x}) \, d\tilde{x}, \quad for \ \xi \in (a, b). \tag{2.2}$$

Proof. The differential equation and boundary condition at $\xi = a$ is satisfied by $\int_{\gamma_a}^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = \mathcal{C}(\xi - a)$, where $\mathcal{C}$ is chosen to satisfy the Dirichlet boundary condition at $\xi = b$. Direct calculation forces $\mathcal{C} = \frac{1}{b-a} \int_{\gamma_a}^{\gamma_b} M(\tilde{x}) \, d\tilde{x}$, and we arrive at the implicit representation (2.2).

We now consider the existence and uniqueness of $x(\xi)$ which satisfies (2.2). The mesh transformation $x(\xi)$ is the solution θ of

$$G(\theta) = \frac{\xi - a}{b - a} \int_{\gamma_a}^{\gamma_b} M(\tilde{x}) \, d\tilde{x}, \tag{2.3}$$

where $G(\theta)$ is defined as $G(\theta) \equiv \int_{\gamma_a}^{\theta} M(\tilde{x}) \, d\tilde{x}$. Under the assumptions of Lemma 2.1, G is continuous. Moreover, G is uniformly monotonic, ie. $\frac{dG}{d\theta} = M(\theta) \ge \check{m} > 0$. Hence, by the implicit function theorem, a unique, continuously differentiable solution to (2.3) and (2.2) results. $\square$

The analysis of the optimized Schwarz methods in Section 5 will require the solution of boundary value problems of the form

$$\frac{d}{d\xi}\left(M(x)\frac{dx}{d\xi}\right) = 0, \quad x(0) = 0, \quad M(x)x_\xi + px\Big|_b = \gamma_b, \tag{2.4}$$

where p and γ_b are constants and $b \in (0, 1)$ is fixed.

LEMMA 2.2. *Under the assumptions of Lemma 2.1 the BVP (2.4) has a unique solution for all $p > 0$ given implicitly by*

$$\int_0^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = (\gamma_b - px(b))\xi, \quad \text{for } \xi \in (0, b). \tag{2.5}$$

Proof. The differential equation and boundary condition at $\xi = 0$ is satisfied by $\int_0^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = \mathcal{C}\xi$, where $\mathcal{C}$ needs to be chosen to satisfy the Robin type boundary condition at $\xi = b$. Direct calculation forces $\mathcal{C} = \gamma_b - px(b)$, from which the implicit representation (2.5) results.

We now consider the existence and uniqueness of $x(\xi)$ which satisfies (2.5). If $\xi = b$ then the boundary value $x(b)$ is the solution θ of

$$G(\theta) = b\gamma_b, \tag{2.6}$$

where $G(\theta)$ is defined as $G(\theta) \equiv \int_0^\theta M(\tilde{x}) \, d\tilde{x} + bp\theta$.

Under the assumptions of Lemma 2.1, G is continuous. Moreover G is uniformly monotonic, ie. there exists a constant $G_p > 0$ such that $\frac{dG}{d\theta} = M(\theta) + bp \geq G_p > 0$. Hence (2.6) has a unique solution $x(b)$. The unique, continuously differentiable solution $x(\xi)$, for $\xi \in (0, b)$, follows by considering (2.5) for the now specified $x(B)$ and noting that the map $\tilde{G}(\theta) = \int_0^\theta M(\tilde{x}) \, d\tilde{x}$ is continuous and uniformly monotonic and hence has a continuously differentiable inverse. $\square$

We will also be interested in solutions of Robin problems of the form

$$\frac{d}{d\xi}\left(M(x)\frac{dx}{d\xi}\right) = 0, \quad M(x)x_\xi - px\Big|_b = \gamma_b, \quad x(1) = 1, \tag{2.7}$$

where p and γ_b are constants and $B \in (0, 1)$ is fixed. Notice the change of sign in the boundary condition at $\xi = b$.

LEMMA 2.3. *Under the assumptions of Lemma 2.1 the BVP (2.7) has a unique solution for all $p > 0$ given implicitly as*

$$\int_{x(\xi)}^1 M(\tilde{x}) \, d\tilde{x} = (\gamma_b + px(b))(1 - \xi), \quad \text{for } \xi \in (b, 1). \tag{2.8}$$

Proof. The representation (2.8) is found by a calculation similar to that in Lemma 2.2. Existence and uniqueness follows as in the previous Lemma by considering the monotonicity of the maps

$$G(\theta) = \int_\theta^1 M(\tilde{x}) \, d\tilde{x} - (1 - b)(\gamma_b + \theta p) \quad \text{and} \quad \tilde{G}(\theta) = \int_\theta^1 M(\tilde{x}) \, d\tilde{x}. \quad \square$$

In Section 7 the time discretized subdomain problems require the solution of nonlinear BVPs of the form

$$x - q\frac{d}{d\xi}\left(M(x)\frac{dx}{d\xi}\right) = f, \quad x(\xi, 0) = x_0(\xi), \quad x(a) = \gamma_a, \quad x(b) = \gamma_b, \tag{2.9}$$

where q is a given constant and $f = f(\xi)$ a given source function. Under the assumptions of Lemma 2.1 the well–posedness of (2.9) can be found, for example, in [26].

3. Parallel Schwarz Methods. In this section, we propose the solution of (1.4) by various classical, parallel Schwarz iterations and consider the convergence of the resulting iterations. Because of the nonlinear nature of (1.4), convergence does not immediately follow from classical analysis of the Schwarz method. In addition, we are interested in precise estimates of the contraction factors of each method.

3.1. A Parallel Nonlinear Schwarz Method. We decompose the domain $\Omega_c = (0,1)$ into two overlapping subdomains $\Omega_1 = (0,\beta)$ and $\Omega_2 = (\alpha, 1)$ with $\alpha < \beta$, and consider for $n = 1, 2, \ldots$ the iteration

$$
\begin{aligned}
(M(x_1^n)x_{1,\xi}^n)_\xi &= 0, \ \ \xi \in \Omega_1, & (M(x_2^n)x_{2,\xi}^n)_\xi &= 0, \ \ \xi \in \Omega_2, \\
x_1^n(0) &= 0, & x_2^n(\alpha) &= x_1^{n-1}(\alpha), \\
x_1^n(\beta) &= x_2^{n-1}(\beta), & x_2^n(1) &= 1.
\end{aligned}
\tag{3.1}
$$

We first construct implicit solutions on the subdomains using Lemma 2.1.

LEMMA 3.1. *Under the assumptions of Lemma 2.1, the subdomain solutions on Ω_1 and Ω_2 of (3.1) are given implicitly by the formulas*

$$
\int_0^{x_1^n(\xi)} M(\tilde{x}) \, d\tilde{x} = \frac{\xi}{\beta} \int_0^{x_2^{n-1}(\beta)} M(\tilde{x}) \, d\tilde{x}
\tag{3.2}
$$

and

$$
\int_{x_2^n(\xi)}^1 M(\tilde{x}) \, d\tilde{x} = \frac{1-\xi}{1-\alpha} \int_{x_1^{n-1}(\alpha)}^1 M(\tilde{x}) \, d\tilde{x}.
\tag{3.3}
$$

Proof. Simply compare the subdomain problems in (3.1) with (2.1) and use the implicit representation of the solution in (2.2). □

Using these representations of the subdomain solutions, we are now in a position to relate $x_{1,2}^n$ to $x_{1,2}^{n-2}$ at the subdomain interfaces and hence prove a precise convergence estimate for the nonlinear parallel Schwarz method (3.1). We will use the infinity norm defined for any function $f : (a,b) \to \mathbb{R}$ by $||f||_\infty := \sup_{x \in (a,b)} |f(x)|$.

THEOREM 3.2. *Under the assumptions of Lemma 2.1 the overlapping ($\beta > \alpha$) parallel Schwarz iteration (3.1) converges for any starting values $x_1^0(\alpha)$, $x_2^0(\beta)$. Moreover, we have the linear convergence estimates*

$$
||x - x_1^{2n+1}||_\infty \le \rho^n \frac{\hat{m}}{\check{m}} |x(\beta) - x_2^0(\beta)|, \quad ||x - x_2^{2n+1}||_\infty \le \rho^n \frac{\hat{m}}{\check{m}} |x(\alpha) - x_2^0(\alpha)|, \tag{3.4}
$$

with contraction factor $\rho := \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha} < 1$.

Proof. Using Lemma 3.1 and defining $C := \int_0^1 M(\tilde{x}) \, d\tilde{x}$, the sequence $x_1^n(\alpha)$ satisfies

$$
\int_0^{x_1^n(\alpha)} M(\tilde{x}) \, d\tilde{x} = \frac{\alpha}{\beta} \int_0^{x_2^{n-1}(\beta)} M(\tilde{x}) \, d\tilde{x} = \frac{\alpha}{\beta} \left(C - \int_{x_2^{n-1}(\beta)}^1 M(\tilde{x}) \, d\tilde{x} \right)
$$

$$
= \frac{\alpha}{\beta} \left(C - \frac{1-\beta}{1-\alpha} \int_{x_1^{n-2}(\alpha)}^1 M(\tilde{x}) \, d\tilde{x} \right) = \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha} \int_0^{x_1^{n-2}(\alpha)} M(\tilde{x}) \, d\tilde{x} + \frac{\alpha}{\beta} \frac{\beta - \alpha}{1-\alpha} C,
\tag{3.5}
$$

where the second and fourth equalities above follow from $\int_0^x M(\tilde{x}) \, d\tilde{x} = C - \int_x^1 M(\tilde{x}) \, d\tilde{x}$, and the third equality follows from (3.3) evaluated at $\xi = \beta$ with n replaced by $n-1$.

Defining the quantity $K_1^n = \int_0^{x_1^n(\alpha)} M(\tilde{x}) \, d\tilde{x}$, relation (3.5) yields the *linear* fixed point iteration

$$K_1^n = \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha} K_1^{n-2} + \frac{\alpha}{\beta} \frac{\beta-\alpha}{1-\alpha} C. \tag{3.6}$$

Since the contraction factor of this iteration $\rho := \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha}$, is strictly less than one, the iteration converges to a limit K_1^* satisfying

$$K_1^* = \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha} K_1^* + \frac{\alpha}{\beta} \frac{\beta-\alpha}{1-\alpha} C \implies K_1^* = \alpha C. \tag{3.7}$$

We have therefore shown $\lim_{n \to \infty} \int_0^{x_1^n(\alpha)} M(\tilde{x}) \, d\tilde{x} = \alpha \int_0^1 M(\tilde{x}) \, d\tilde{x}$. In order to prove convergence to the correct limit, we note the monodomain solution x also satisfies $\alpha \int_0^1 M(\tilde{x}) \, d\tilde{x} = \int_0^{x(\alpha)} M(\tilde{x}) \, d\tilde{x}$. Hence we have convergence to the correct limit, i.e.

$$\lim_{n \to \infty} \int_0^{x_1^n(\alpha)} M(\tilde{x}) \, d\tilde{x} = \int_0^{x(\alpha)} M(\tilde{x}) \, d\tilde{x}.$$

It remains to prove the convergence estimate in the L^∞ norm. Subtracting equation (3.6) from (3.7), we obtain by induction

$$\int_{x_1^{2n}(\alpha)}^{x(\alpha)} M(\tilde{x}) \, d\tilde{x} = \rho^n \int_{x_1^0(\alpha)}^{x(\alpha)} M(\tilde{x}) \, d\tilde{x}. \tag{3.8}$$

For any $a, b \in \mathbb{R}$ we have, by the boundedness of M,

$$\check{m} |a-b| \leq \left| \int_a^b M(\tilde{x}) \, d\tilde{x} \right| \leq \hat{m} |a-b|. \tag{3.9}$$

Subtracting (3.3) from the equivalent expression for the exact, monodomain solution $x(\xi)$ we obtain

$$\int_{x_2^{2n+1}(\xi)}^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = \frac{1-\xi}{1-\alpha} \int_{x_1^{2n}(\alpha)}^{x(\alpha)} M(\tilde{x}) \, d\tilde{x}.$$

Using (3.8) we have

$$\int_{x_2^{2n+1}(\xi)}^{x(\xi)} M(\tilde{x}) \, d\tilde{x} = \frac{1-\xi}{1-\alpha} \rho^n \int_{x_1^0(\alpha)}^{x(\alpha)} M(\tilde{x}) \, d\tilde{x}.$$

Convergence in the interior is obtained by taking the modulus and using the boundedness of M (3.9). For all $\xi \in [\alpha, 1]$

$$|x(\xi) - x_2^{2n+1}(\xi)| \leq \frac{1-\xi}{1-\alpha} \rho^n \frac{\hat{m}}{\check{m}} |x(\alpha) - x_1^0(\alpha)|.$$

Taking the supremum gives the second estimate in (3.4). The estimate on subdomain one is obtained similarly. $\square$

3.2. A Parallel Linearized Schwarz Method. We may avoid the nonlinear solves on each subdomain in (3.1) by replacing x_1^n and x_2^n in the arguments of the nonlinear function M by the functions obtained from the previous iteration. Consider a linearized, parallel Schwarz iteration: for $n = 1, 2, \ldots$

$$(M(x_1^{n-1})x_{1,\xi}^n)_\xi = 0, \quad \xi \in \Omega_1, \qquad (M(x_2^{n-1})x_{2,\xi}^n)_\xi = 0, \quad \xi \in \Omega_2,$$
$$x_1^n(0) = 0, \qquad\qquad\qquad\qquad x_2^n(\alpha) = x_1^{n-1}(\alpha), \qquad (3.10)$$
$$x_1^n(\beta) = x_2^{n-1}(\beta), \qquad\qquad\qquad x_2^n(1) = 1.$$

Each Schwarz iteration now involves the solution of a linear BVP on each subdomain. A simple calculation yields the following representation of the subdomain solutions.

LEMMA 3.3. *The subdomain solutions of (3.10) are unique and given by*

$$x_1^n(\xi) = x_2^{n-1}(\beta) \frac{\int_0^\xi \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}{\int_0^\beta \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}, \qquad (3.11)$$

and

$$x_2^n(\xi) = x_1^{n-1}(\alpha) + (1 - x_1^{n-1}(\alpha)) \frac{\int_\alpha^\xi \frac{d\tilde{\xi}}{M(x_2^{n-1}(\tilde{\xi}))}}{\int_\alpha^1 \frac{d\tilde{\xi}}{M(x_2^{n-1}(\tilde{\xi}))}}. \qquad (3.12)$$

THEOREM 3.4. *Under the assumptions of Lemma 2.1, the overlapping ($\beta > \alpha$) linearized, parallel Schwarz iteration (3.10) converges for any continuously differentiable initial guesses $x_1^0(\xi)$ and $x_2^0(\xi)$.*

Proof. We first note that in this linearized iteration, one gains regularity: if the initial guess is continuously differentiable, the explicit solution formulas (3.11) and (3.12) show that after one iteration the iterates are already twice continuously differentiable, and hence the algorithm produces a unique sequence of classical solutions. We now demonstrate convergence on subdomain one only, the argument on subdomain two is similar. Evaluating $x_2^{n-1}(\beta)$ from (3.12) and substituting into (3.11) we have, for any $\xi \in (0, \beta]$

$$x_1^n(\xi) = \mathcal{C}_\xi^n x_1^{n-2}(\alpha) + \mathcal{D}_\xi^n, \qquad (3.13)$$

where

$$\mathcal{C}_\xi^n = \frac{\int_\beta^1 \frac{d\tilde{\xi}}{M(x_2^{n-2}(\tilde{\xi}))}}{\int_\alpha^1 \frac{d\tilde{\xi}}{M(x_2^{n-2}(\tilde{\xi}))}} \frac{\int_0^\xi \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}{\int_0^\beta \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}, \qquad \mathcal{D}_\xi^n = \frac{\int_\alpha^\beta \frac{d\tilde{\xi}}{M(x_2^{n-2}(\tilde{\xi}))}}{\int_\alpha^1 \frac{d\tilde{\xi}}{M(x_2^{n-2}(\tilde{\xi}))}} \frac{\int_0^\xi \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}{\int_0^\beta \frac{d\tilde{\xi}}{M(x_1^{n-1}(\tilde{\xi}))}}.$$

The quantities $\mathcal{C}_\xi^n$ and $\mathcal{D}_\xi^n$ satisfy $0 < \mathcal{C}_\xi^n \le \rho_\xi < 1$ and $0 < \mathcal{D}_\xi^n \le \gamma_\xi < 1$, where

$$\rho_\xi := \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{\beta - \alpha}{1 - \beta}} \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{\beta - \xi}{\xi}} \quad \text{and} \quad \gamma_\xi := \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{\beta - \xi}{\xi}} \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{1 - \beta}{\beta - \alpha}}.$$

Furthermore, these quantities are uniformly bounded:

$$\rho_\xi \le \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{\beta - \alpha}{1 - \beta}} := \tilde{\rho} < 1 \quad \text{and} \quad \gamma_\xi \le \frac{1}{1 + \frac{\check{m}}{\hat{m}} \frac{1 - \beta}{\beta - \alpha}} := \tilde{\gamma} < 1.$$

8

To establish these bounds let $F(x) := 1/M(x)$. The assumptions on M imply $\frac{1}{\hat{m}} \leq F(x) \leq \frac{1}{\check{m}}$. As an example, the upper and lower bounds on F then imply

$$\frac{\int_0^\xi F(x(\tilde{\xi}))\,d\tilde{\xi}}{\int_0^\beta F(x(\tilde{\xi}))\,d\tilde{\xi}} = \frac{\int_0^\xi F(x(\tilde{\xi}))\,d\tilde{\xi}}{\int_0^\xi F(x(\tilde{\xi}))\,d\tilde{\xi} + \int_\xi^\beta F(x(\tilde{\xi}))\,d\tilde{\xi}} = \frac{1}{1 + \frac{\int_\xi^\beta F(x(\tilde{\xi}))\,d\tilde{\xi}}{\int_0^\xi F(x(\tilde{\xi}))\,d\tilde{\xi}}} \leq \frac{1}{1 + \frac{\check{m}}{\hat{m}}\frac{\beta-\xi}{\xi}}.$$

Using (3.13) recursively, assuming n is even, we have

$$x_1^n(\xi) = \mathcal{C}_\xi^n \prod_{k=1}^{\frac{n-2}{2}} \mathcal{C}_\alpha^{2k} x_1^0(\alpha) + \mathcal{D}_\xi^n + \mathcal{C}_\xi^n \sum_{k=1}^{\frac{n-2}{2}} \mathcal{D}_\alpha^{2k} \left(\prod_{l=k+1}^{\frac{n-2}{2}} \mathcal{C}_\alpha^{2l} \right),$$

where the product in the k–th term of the sum is assumed to be one if the lower index of the product exceeds the upper index. Since $\tilde{\rho} < 1$, the product multiplying $x_1^0(\alpha)$ must go to zero as $n \to \infty$. The infinite series converges by direct comparison with $\sum_{k=1}^\infty \tilde{\gamma}\tilde{\rho}^{k-1}$. Indeed by the Weierstrass M–test [3], $\{x_1^n(\xi)\}$ converges uniformly to some limiting function $\tilde{x}_1$. Likewise $\{x_2^n(\xi)\}$ converges uniformly to some limiting function $\tilde{x}_2$. Since $\tilde{x}_1(\alpha) = \tilde{x}_2(\alpha)$ and $\tilde{x}_1(\beta) = \tilde{x}_2(\beta)$, both $\tilde{x}_1$ and $\tilde{x}_2$ satisfy the same PDE in the overlap with the same two boundary conditions, and by uniqueness, $\tilde{x}_1$ and $\tilde{x}_2$ must coincide in the overlap. One can therefore simply glue these two solutions together in order to obtain a function which satisfies the PDE in all of Ω_c, and also the two original boundary conditions at 0 and 1. Again by uniqueness, this must now be the desired solution. □

Alternating Schwarz versions of the algorithms presented in Sections 3.1 and 3.2 are given in [22]. These alternating algorithms obtain faster convergence by using the most recently computed boundary information at the subdomain interfaces. As an example, the second transmission condition in (3.10) would be changed to $x_2^n(\alpha) = x_1^n(\alpha)$. The improved convergence is obtained at the cost of losing obvious parallelization.

4. An Optimal Schwarz Method. The classical Schwarz methods in Section 3 converge slowly — linear convergence with a convergence rate depending on the size of the overlap. Increasing the overlap improves the rate of contraction but at a cost of more expensive subdomain solves. Moreover, in many physical situations it is not possible to allow subdomains to overlap. Without overlap the classical Schwarz methods developed in Section 3 would not converge.

Another strategy to improve convergence, and possibly avoid overlap altogether, is to consider alternate transmission conditions at the subdomain interfaces. In this section we consider non–overlapping *optimal Schwarz* methods which guarantee convergence in two iterations on two subdomains. These methods can also be used with overlap (and will be presented elsewhere), but this does not change their performance — even with overlap convergence can not be achieved in less than two iterations.

We consider the development of optimal Schwarz methods for boundary value problems of the form (1.4). We decompose $\Omega_c = (0,1)$ into two non–overlapping subdomains $\Omega_1 = (0,\alpha)$ and $\Omega_2 = (\alpha,1)$ and consider the iteration

$$\begin{aligned}
(M(x_1^n)x_{1,\xi}^n)_\xi &= 0, \quad \xi \in \Omega_1, & (M(x_2^n)x_{2,\xi}^n)_\xi &= 0, \quad \xi \in \Omega_2, \\
x_1^n(0) &= 0, & \mathcal{B}_2(x_2^n(\alpha)) &= \mathcal{B}_2(x_1^{n-1}(\alpha)), & (4.1) \\
\mathcal{B}_1(x_1^n(\alpha)) &= \mathcal{B}_1(x_2^{n-1}(\alpha)), & x_2^n(1) &= 1,
\end{aligned}$$

where the transmission operators $\mathcal{B}_1$ and $\mathcal{B}_2$ are given by

$$\mathcal{B}_1(\cdot) \equiv M(\cdot)\partial_\xi(\cdot) - \mathcal{S}_1(\cdot), \quad \mathcal{B}_2(\cdot) \equiv M(\cdot)\partial_\xi(\cdot) - \mathcal{S}_2(\cdot) \tag{4.2}$$

and

$$\mathcal{S}_1(\cdot) = \frac{1}{1-\alpha} \int_{(\cdot)}^{1} M(\tilde{x})\, d\tilde{x}, \quad \text{and} \quad \mathcal{S}_2(\cdot) = \frac{1}{\alpha} \int_{0}^{(\cdot)} M(\tilde{x})\, d\tilde{x}. \tag{4.3}$$

The subdomain solutions are given by (2.2) where $x_1^n(\alpha)$ and $x_2^n(\alpha)$ are determined by the transmission conditions.

THEOREM 4.1. *Under the assumptions of Lemma 2.1, the iteration (4.1) with transmission conditions specified by (4.2–4.3) is optimal, ie. convergence is achieved in two iterations.*

Proof. The derivatives of each subdomain solution satisfy

$$M(x_1^n(\xi))\partial_\xi x_1^n(\xi) = \frac{1}{\alpha} \int_{0}^{x_1^n(\alpha)} M(\tilde{x})\, d\tilde{x} \quad \text{and} \quad M(x_2^n(\xi))\partial_\xi x_2^n(\xi) = \frac{1}{1-\alpha} \int_{x_2^n(\alpha)}^{1} M(\tilde{x})\, d\tilde{x}.$$

Suppose $x_1^{n-1}(\xi)$ and $x_2^{n-1}(\xi)$ are given iterates on subdomain one and two. The transmission condition on subdomain one requires

$$\frac{1}{\alpha} \int_{0}^{x_1^n(\alpha)} M(\tilde{x})\, d\tilde{x} - \frac{1}{1-\alpha} \int_{x_1^n(\alpha)}^{1} M(\tilde{x})\, d\tilde{x}$$

$$= \frac{1}{1-\alpha} \int_{x_2^{n-1}(\alpha)}^{1} M(\tilde{x})\, d\tilde{x} - \frac{1}{1-\alpha} \int_{x_2^{n-1}(\alpha)}^{1} M(\tilde{x})\, d\tilde{x} = 0,$$

which implies

$$\frac{1}{\alpha} \int_{0}^{x_1^n(\alpha)} M(\tilde{x})\, d\tilde{x} - \frac{1}{1-\alpha} \int_{x_1^n(\alpha)}^{1} M(\tilde{x})\, d\tilde{x} = 0.$$

Using the fact that $\int_{x_1^n(\alpha)}^{1} M(\tilde{x})\, d\tilde{x} = \int_{0}^{1} M(\tilde{x})\, d\tilde{x} - \int_{0}^{x_1^n(\alpha)} M(\tilde{x})\, d\tilde{x}$, we have

$$\int_{0}^{x_1^n(\alpha)} M(\tilde{x})\, d\tilde{x} = \alpha \int_{0}^{1} M(\tilde{x})\, d\tilde{x}.$$

As mentioned previously, the exact solution also satisfies $\int_{0}^{x(\alpha)} M(\tilde{x})\, d\tilde{x} = \alpha \int_{0}^{1} M(\tilde{x})\, d\tilde{x}$. Since M is strictly positive and integrable then $x_1^n(\alpha) = x(\alpha)$, that is we have convergence in two iterations at the boundary. Convergence in the interior of the domain follows from well–posedness of the BVP. A similar calculation holds on subdomain two. $\square$

5. Optimized Schwarz Methods. As mentioned, the slow convergence of classical Schwarz can often be improved by using transmission conditions which are different from the classical Dirichlet conditions at the subdomain interfaces. The optimal conditions derived in Section 4 are non–local and are expensive to use: they require the evaluation of an integral. In this section we approximate the optimal conditions by developing a Robin type transmission condition. We will only consider the non–overlapping case, the overlapping case can be handled in a similar manner and will be presented elsewhere.

On subdomain one the optimal transmission condition is

$$M(x_1^n(\alpha))\partial_\xi x_1^n(\alpha) - \frac{\int_{x_1^n(\alpha)}^1 M(\tilde{x})\,d\tilde{x}}{1-\alpha} = M(x_2^{n-1}(\alpha))\partial_\xi x_2^{n-1}(\alpha) - \frac{\int_{x_2^{n-1}(\alpha)}^1 M(\tilde{x})\,d\tilde{x}}{1-\alpha}.$$

Using $\int_x^1 M(\tilde{x})\,d\tilde{x} = C - \int_0^x M(\tilde{x})\,d\tilde{x}$, where, as before, $C = \int_0^1 M(\tilde{x})\,d\tilde{x}$, the boundary condition may be rewritten as

$$M(x_1^n(\alpha))\partial_\xi x_1^n(\alpha) - \frac{C - \int_0^{x_1^n(\alpha)} M(\tilde{x})\,d\tilde{x}}{1-\alpha} = M(x_2^{n-1}(\alpha))\partial_\xi x_2^{n-1}(\alpha) - \frac{C - \int_0^{x_2^{n-1}(\alpha)} M(\tilde{x})\,d\tilde{x}}{1-\alpha}.$$

Canceling the $(1-\alpha)^{-1}C$ terms and using the mean value theorem for integrals we can rewrite this expression as

$$M(x_1^n(\alpha))\partial_\xi x_1^n(\alpha) + \frac{1}{1-\alpha}M(x_1^*)x_1^n(\alpha) = M(x_2^{n-1}(\alpha))\partial_\xi x_2^{n-1}(\alpha) + \frac{1}{1-\alpha}M(x_2^*)x_2^{n-1}(\alpha),$$

or

$$M(x_1^n(\alpha))\partial_\xi x_1^n(\alpha) + \bar{p}_1^n x_1^n(\alpha) = M(x_2^{n-1}(\alpha))\partial_\xi x_2^{n-1}(\alpha) + \hat{p}_1^n x_2^{n-1}(\alpha), \qquad (5.1)$$

where $\bar{p}_1^n, \hat{p}_1^n > 0$ (since M is strictly positive).

Likewise, on subdomain two we can obtain

$$M(x_2^n(\alpha))\partial_\xi x_2^n(\alpha) - \bar{p}_2^n x_2^n(\alpha) = M(x_1^{n-1}(\alpha))\partial_\xi x_1^{n-1}(\alpha) - \hat{p}_2^n x_1^{n-1}(\alpha), \qquad (5.2)$$

where $\bar{p}_2^n, \hat{p}_2^n > 0$.

This suggests approximating the optimal transmission conditions (5.1) and (5.2) as in the following algorithm. Decompose $\Omega_c = (0,1)$ into two non–overlapping sub-domains $\Omega_1 = (0,\alpha)$ and $\Omega_2 = (\alpha,1)$, and consider the iteration: for $n = 1, 2, \dots$

$$
\begin{aligned}
(M(x_1^n)x_{1,\xi}^n)_\xi = 0, \quad \xi \in \Omega_1, && (M(x_2^n)x_{2,\xi}^n)_\xi = 0, \quad \xi \in \Omega_2, \\
x_1^n(0) = 0, && \widetilde{\mathcal{B}}_2(x_2^n(\alpha)) = \widetilde{\mathcal{B}}_2(x_1^{n-1}(\alpha)), \\
\widetilde{\mathcal{B}}_1(x_1^n(\alpha)) = \widetilde{\mathcal{B}}_1(x_2^{n-1}(\alpha)) && x_2^n(1) = 1, \qquad (5.3)
\end{aligned}
$$

where the transmission operators $\widetilde{\mathcal{B}}_i$, $i = 1, 2$ are given by

$$\widetilde{\mathcal{B}}_1(\cdot) \equiv M(\cdot)\partial_\xi(\cdot) + pI(\cdot), \quad \widetilde{\mathcal{B}}_2(\cdot) \equiv M(\cdot)\partial_\xi(\cdot) - pI(\cdot),$$

with $I(\cdot)$ the identity operator and p a constant to be chosen to improve the convergence rate. By comparing with the optimal transmission conditions given in (4.2), we see that if one knew the values of the integrals computed by the operators $\mathcal{S}_i$, $i = 1, 2$ for a given situation, one could actually choose four corresponding values and obtain an optimal algorithm in this one dimensional setting. In general however, especially in higher dimensions, such a precise choice is not practical, and we will see in Section 8 that already with one parameter p one can obtain a substantially improved algorithm.

We now analyze the optimized Schwarz algorithm (5.3). The well–posedness of the subdomain problems was established in Lemma 2.2 and 2.3. The subdomain solutions are represented implicitly in the following Lemma.

LEMMA 5.1. *Under the assumptions of Lemmas 2.2 and 2.3, the subdomain solutions on Ω_1 and Ω_2 of (5.3) are given implicitly by the formulas*

$$\int_0^{x_1^n(\xi)} M(\tilde{x})\,d\tilde{x} = R_1(x_1^n(\alpha))\xi \quad and \quad \int_{x_2^n(\xi)}^1 M(\tilde{x})\,d\tilde{x} = R_2(x_2^n(\alpha))(1-\xi),$$

where the operators R_1 and R_2 are given by

$$R_1(x) = \frac{1}{\alpha} \int_0^x M(\tilde{x})\, d\tilde{x} \quad and \quad R_2(x) = \frac{1}{1-\alpha} \int_x^1 M(\tilde{x})\, d\tilde{x}. \qquad (5.4)$$

The Robin conditions at the interface force the operator values to satisfy the recurrence relations:

$$R_1(x_1^{n+1}(\alpha)) + px_1^{n+1}(\alpha) = R_2(x_2^n(\alpha)) + px_2^n(\alpha) \qquad (5.5)$$

and

$$R_2(x_2^n(\alpha)) - px_2^n(\alpha) = R_1(x_1^{n-1}(\alpha)) - px_1^{n-1}(\alpha). \qquad (5.6)$$

Proof. The representation of the subdomain solutions follows from Lemmas 2.2 and 2.3. The recurrence relations (5.5) and (5.6) are a consequence of the transmission conditions at $\xi = \alpha$. $\square$

The iteration (5.5–5.6) is of the Peaceman–Rachford type [42], and our convergence proof follows the discussion of nonlinear Peaceman–Rachford iterations in [34, 41].

THEOREM 5.2. *Under the assumptions of Lemma 2.1 the iteration (5.5–5.6) converges globally to the exact solution $x(\alpha)$ for all $p > 0$. Moreover, we have the linear convergence estimate*

$$||x - x_1^{2n+1}||_\infty \leq \frac{\hat{m}}{\check{m}} \cdot \frac{p + \frac{1}{\alpha}\hat{m}}{p + \frac{1}{\alpha}\check{m}} \rho_{robin}^n |x(\alpha) - x_2^0(\alpha)|, \quad ||x - x_2^{2n+1}||_\infty \leq \frac{\hat{m}}{\check{m}} \cdot \frac{p + \frac{1}{1-\alpha}\hat{m}}{p + \frac{1}{1-\alpha}\check{m}} \rho_{robin}^n |x(\alpha) - x_2^0(\alpha)|,$$

where an estimate on the contraction factor is

$$\rho_{robin} = \sqrt{\frac{p^2 + \frac{\hat{m}^2}{(1-\alpha)^2} - 2p\frac{\check{m}}{1-\alpha}}{p^2 + \frac{\hat{m}^2}{(1-\alpha)^2} + 2p\frac{\check{m}}{1-\alpha}}} \cdot \sqrt{\frac{p^2 + \frac{\hat{m}^2}{\alpha^2} - 2p\frac{\check{m}}{\alpha}}{p^2 + \frac{\hat{m}^2}{\alpha^2} + 2p\frac{\check{m}}{\alpha}}}. \qquad (5.7)$$

Proof. We rewrite (5.5) and (5.6) as

$$(pI - R_2)x_2^n(\alpha) = (pI - R_1)x_1^{n-1}(\alpha), \ and \ (pI + R_1)x_1^{n+1}(\alpha) = (pI + R_2)x_2^n(\alpha).$$

The operators R_1 and $-R_2$ defined in (5.4) are continuous and uniformly monotonic (increasing) since

$$R_1'(x) = \frac{1}{\alpha}M(x) \geq \frac{1}{\alpha}\check{m} > 0 \quad and \quad -R_2'(x) = \frac{1}{1-\alpha}M(x) \geq \frac{1}{1-\alpha}\check{m} > 0.$$

Since $p > 0$, $pI - R_2$ and $pI + R_1$ are also continuous and uniformly monotonic and hence invertible. This implies that $x_2^n(\alpha)$ and $x_1^{n+1}(\alpha)$ are well–defined. Eliminating $x_2^n(\alpha)$ we obtain the recursion $x_1^{n+1}(\alpha) = Gx_1^{n-1}(\alpha)$, where

$$G \equiv (pI + R_1)^{-1}(pI + R_2)(pI - R_2)^{-1}(pI - R_1).$$

The operators R_1 and $-R_2$ are uniformly monotone and Lipschitz, thus

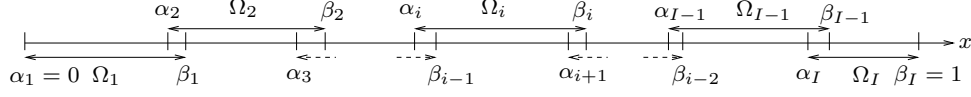$$G_1 = (pI + R_2)(pI - R_2)^{-1} \quad and \quad G_2 = (pI - R_1)(pI + R_1)^{-1}$$

12

FIG. 6.1. *Decomposition into many subdomains.*

are strict contractions for all $p > 0$, see p. 388 of [41] for details. Therefore, the iteration $z^n(\alpha) = G_1 G_2 z^{n-2}(\alpha)$, with $z^0(\alpha) = (pI + R_1) x_1^0(\alpha)$, will converge. Since $G = (pI + R_1)^{-1} G_1 G_2 (pI + R_1)$ and $z^{2n}(\alpha) = (pI + R_1) x_1^{2n}(\alpha)$ then $x_1^{2n}(\alpha)$ will also converge for any $x_1^0(\alpha)$ to some limit $x_1^*(\alpha)$. A similar argument is possible to show the odd iterates $x_1^{2n+1}(\alpha)$ converge to the same limit. Likewise, the sequence $x_2^n(\alpha)$ converges to a limit point $x_2^*(\alpha)$. The points $x_1^*(\alpha)$ and $x_2^*(\alpha)$ must satisfy (5.5) and (5.6) in the limit. Adding the limits of (5.5) and (5.6) we obtain $x_1^*(\alpha) = x_2^*(\alpha) =: x^*(\alpha)$. And taking the difference we find the limit point $x^*(\alpha)$ satisfies $R_1(x^*(\alpha)) = R_2(x^*(\alpha))$ or

$$\frac{1}{\alpha} \int_0^{x^*(\alpha)} M(\tilde{x})\, d\tilde{x} = \frac{1}{\alpha - 1} \left(\int_0^{x^*(\alpha)} M(\tilde{x})\, d\tilde{x} - C \right).$$

An argument similar to that in Theorems 3.2 and 4.1 shows that $x^*(\alpha) = x(\alpha)$.

The contraction factor, ρ_{robin}, for the sequence $z^n(\alpha)$, given by (5.7), is found by computing the Lipschitz constant of the operator $G_1 G_2$ as the product of the Lipschitz constants of G_1 and G_2, see [41] for details. The convergence factor of $x_1^n(\alpha)$ is related to ρ_{robin}. In fact,

$$|x^*(\alpha) - x_1^{2n}(\alpha)| \le L|z^*(\alpha) - z^{2n}(\alpha)| \le L\rho_{\text{robin}}^n |z^*(\alpha) - z^0(\alpha)| \le L\tilde{L}\rho_{\text{robin}}^n |x^*(\alpha) - x_1^0(\alpha)|,$$

where L and $\tilde{L}$ are the Lipschitz constants for $(pI + R_1)^{-1}$ and $(pI + R_1)$ respectively. Inspection of these two operators yields $L = (p + \frac{1}{\alpha}\check{m})^{-1}$ and $\tilde{L} = p + \frac{1}{\alpha}\hat{m}$. Together with the familiar interior estimate $|x_1^{2n}(\xi) - x(\xi)| \le \frac{\hat{m}}{\check{m}}|x(\alpha) - x_1^{2n}(\alpha)|$, we have

$$|x_1^{2n}(\xi) - x(\xi)| \le \frac{\hat{m}}{\check{m}} \cdot \frac{p + \frac{1}{\alpha}\hat{m}}{p + \frac{1}{\alpha}\check{m}} \rho_{\text{robin}}^n |x(\alpha) - x_1^0(\alpha)|.$$

A similar analysis gives the interior estimate on subdomain two. □

6. Multidomains. In practice, we would like to take advantage of today's multicore, multiprocessor environments for grid generation. To this end, we now consider the extension of the parallel, nonlinear, classical Schwarz algorithm presented in Section 3.1, to $I > 2$ subdomains, see Figure 6.1. The solution, $x_i(\xi)$, on the i–th subdomain, $\Omega_i = (\alpha_i L, \beta_i L)$, $i = 1, 2, \ldots, I$ is found by solving

$$(M(x_i)x_{i,\xi})_\xi = 0, \quad x_i(\alpha_i) = x_{i-1}(\alpha_i), \quad x_i(\beta_i) = x_{i+1}(\beta_i), \tag{6.1}$$

where $\alpha_1 = 0$, $x_0(\alpha_1) = 0$, $\beta_I = 1$ and $x_{I+1}(\beta_I) = 1$. We assume that $\beta_i \le \alpha_{i+2}$ for $i = 1, \ldots, J - 2$ so that there is no overlap between non–adjacent subdomains. We obtain the solution on the whole domain by composing the subdomain solutions $x_i(\xi)$.

The subdomain solutions $x_i(\xi)$ are found by the classical parallel Schwarz iteration: for $n = 1, 2, \ldots$

$$(M(x_i^n)x_{i,\xi}^n)_\xi = 0, \quad x_i^n(\alpha_i) = x_{i-1}^{n-1}(\alpha_i), \quad x_i^n(\beta_i) = x_{i+1}^{n-1}(\beta_i), \tag{6.2}$$

13

for $i = 1, \ldots, I$, where we have defined for convenience, $x_0^n(\alpha_1) \equiv 0$ and $x_{I+1}^n(\beta_I) \equiv 1$.

Our proof follows the general idea given in [16] for the Schwarz waveform relaxation algorithm applied to the linear heat equation, but with a couple of important alterations due to the nonlinearity of the problem (6.1). Furthermore, we provide an alternative, simpler analysis of the linear matrix inequalities required to obtain the contraction.

We define the error on the ith subdomain, at iteration n, as

$$e_i^n(\xi) = \int_{x_i(\xi)}^{x_i^n(\xi)} M(\tilde{x}) \, d\tilde{x}, \qquad \text{for} \quad i = 1, \ldots, I. \tag{6.3}$$

Convergence is demonstrated by showing this measure of the error contracts to zero on all subdomains. As in Section 3.1, it is easy to see that if M is bounded away from zero then $\lim_{n \to \infty} e_i^n(\xi) = 0$ implies $\lim_{n \to \infty} |x_i(\xi) - x_i^n(\xi)| = 0$.

The error on each subdomain is given explicitly in the following Lemma. For notational convenience we introduce $e_0^n(\alpha_1) \equiv 0$ and $e_{I+1}^n(\beta_I) \equiv 0$.

LEMMA 6.1. *The error on each subdomain satisfies*

$$e_i^{n+2}(\xi) = \frac{1}{\beta_i - \alpha_i} \left[(\xi - \alpha_i) e_{i+1}^{n+1}(\beta_i) + (\beta_i - \xi) e_{i-1}^{n+1}(\alpha_i) \right], \quad for \quad \xi \in [\alpha_i, \beta_i], \tag{6.4}$$

and for each $i = 1, \ldots, I$.

Proof. Subtracting (6.1) from (6.2) and differentiating the error expression (6.3) twice, we see the error $e_i^{n+2}(\xi)$ satisfies the linear BVP

$$\frac{d^2 e_i^{n+2}}{d\xi^2} = 0, \quad e_i^{n+2}(\alpha_i) = e_{i-1}^{n+1}(\alpha_i), \quad e_i^{n+2}(\beta_i) = e_{i+1}^{n+1}(\beta_i),$$

for $i = 1, \ldots, I$. Direct integration and enforcing the boundary conditions gives the result. □

As in the two subdomain case, we now use (6.4) to relate the error on subdomain i at iteration $n + 2$ to the error on subdomain i and its neighbors at iteration n. Following [16], we introduce the following quantities:

$$r_i = \frac{\beta_{i-1} - \alpha_i}{\beta_i - \alpha_i}, \quad p_i = \frac{\beta_i - \beta_{i-1}}{\beta_i - \alpha_i}, \quad q_i = \frac{\alpha_{i+1} - \alpha_i}{\beta_i - \alpha_i}, \quad \text{and} \quad s_i = \frac{\beta_i - \alpha_{i+1}}{\beta_i - \alpha_i}. \tag{6.5}$$

LEMMA 6.2. *The error at the interface $\xi = \beta_{i-1}$, $i = 2, \ldots, N$ satisfies*

$$|e_i^{n+2}(\beta_{i-1})| \leq r_i r_{i+1} |e_{i+2}^n(\beta_{i+1})| + r_i p_{i+1} |e_i^n(\alpha_{i+1})| + p_i q_{i-1} |e_i^n(\beta_{i-1})| + p_i s_{i-1} |e_{i-2}^n(\alpha_{i-1})|, \tag{6.6}$$

while at $\xi = \alpha_{i+1}$, $i = 1, \ldots, N - 1$ we have

$$|e_i^{n+2}(\alpha_{i+1})| \leq q_i r_{i+1} |e_{i+2}^n(\beta_{i+1})| + q_i p_{i+1} |e_i^n(\alpha_{i+1})| + s_i q_{i-1} |e_i^n(\beta_{i-1})| + s_i s_{i-1} |e_{i-2}^n(\alpha_{i-1})|. \tag{6.7}$$

Proof. Inequality (6.6) is obtained by evaluating (6.4) at β_{i-1} and once again using (6.4) twice more to write $e_{i+1}^{n+1}(\beta_i)$ in terms of e_{i+2}^n and e_i^n and $e_{i-1}^{n+1}(\alpha_i)$ in terms of e_i^n and e_{i-2}^n. Taking absolute values, using the triangle inequality and noting r_i, p_i, q_i and s_i are non-negative gives the result. Inequality (6.7) is obtained in a similar way. □

Assuming an even number of subdomains (the odd case can be handled in a similar manner), the relations in (6.2) may be written as matrix inequalities

$$\boldsymbol{e}^{n+2} \le M_{\boldsymbol{e}} \boldsymbol{e}^n \quad \text{and} \quad \hat{\boldsymbol{e}}^{n+2} \le M_{\hat{\boldsymbol{e}}} \hat{\boldsymbol{e}}^n, \tag{6.8}$$

where

$$\boldsymbol{e}^n = (|e_1^n(\alpha_2)|, |e_3^n(\beta_2)|, |e_3^n(\alpha_4)|, \ldots, |e_{I-1}^n(\beta_{I-2})|, |e_{I-1}^n(\alpha_I)|)^T,$$
$$\hat{\boldsymbol{e}}^n = (|e_2(\beta_1)|, |e_2^n(\alpha_3)|, |e_4^n(\beta_3)|, \ldots, |e_{I-2}^n(\alpha_{I-1})|, |e_I^n(\beta_{I-1})|)^T,$$

and the $(I-1) \times (I-1)$ matrices $M_{\boldsymbol{e}}$ and $M_{\hat{\boldsymbol{e}}}$ are given as

$$M_{\boldsymbol{e}} = \begin{pmatrix} q_1 p_2 & q_1 r_2 & & & & & & \\ p_3 s_2 & p_3 q_2 & r_3 p_4 & r_3 r_4 & & & & \\ s_3 s_2 & s_3 q_2 & q_3 p_4 & q_3 r_4 & & & & \\ & & p_5 s_4 & p_5 q_4 & r_5 p_6 & & r_5 r_6 & \\ & & & \ddots & & \ddots & & \\ & & & & p_{I-1} s_{I-2} & p_{I-1} q_{I-2} & r_{I-1} p_I \\ & & & & s_{I-1} s_{I-2} & s_{I-1} q_{I-2} & q_{I-1} p_I \end{pmatrix},$$

and

$$M_{\hat{\boldsymbol{e}}} = \begin{pmatrix} p_2 q_1 & r_2 q_3 & r_2 r_3 & & & & & \\ s_2 q_1 & q_2 p_3 & q_2 r_3 & & & & & \\ & p_4 s_3 & p_4 q_3 & r_4 p_5 & & r_4 r_5 & & \\ & s_4 s_3 & s_4 q_3 & q_4 p_5 & & q_4 r_5 & & \\ & & & \ddots & & \ddots & & \\ & & & p_{I-2} s_{I-3} & p_{I-2} q_{I-3} & r_{I-2} p_{I-1} & r_{I-2} r_{I-1} \\ & & & s_{I-2} s_{I-3} & s_{I-2} q_{I-3} & q_{I-2} p_{I-1} & q_{I-2} r_{I-1} \\ & & & & & p_I s_{I-1} & p_I q_{I-1} \end{pmatrix}.$$

Inspecting (6.5), we see that $\|M_{\boldsymbol{e}}\|_\infty = \|M_{\hat{\boldsymbol{e}}}\|_\infty = 1$; hence it is not obvious that the inequalities (6.8) converge to zero as $n \to \infty$. In [16], the authors use a delicate analysis to show that there exists some $\tilde{n} > 0$ for which $\|M_{\boldsymbol{e}}^{\tilde{n}}\|_\infty < 1$ and $\|M_{\hat{\boldsymbol{e}}}^{\tilde{n}}\|_\infty < 1$ from which convergence follows. Here we demonstrate another technique using the recent work [49]. Suppose A is a real $m \times m$ matrix and a a positive real number. Then $\Psi_a(A)$ is the matrix obtained by deleting all rows, and corresponding columns, for which the row sum $\sum_{j=1}^m |a_{ij}| < a$. For example, if

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix} \implies \Psi_1(A) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{3}{4} \end{pmatrix} \text{ and } \Psi_1^2(A) = \Psi_1(\Psi_1(A)) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The following result, which we state as a Lemma, was given in [49].

LEMMA 6.3. *Suppose a $m \times m$ real matrix A is element–wise non–negative, then $\rho(A) < \|A\|_\infty$ if and only if $\Psi_{\|A\|_\infty}^m(A) = 0$. Here 0 is the $m \times m$ zero matrix.*

For the matrix A above, $\|A\|_\infty = 1$ and $\Psi_1^2(A) = 0$ hence Lemma 6.3 guarantees $\rho(A) < 1$. Direct calculation of the eigenvalues shows $\rho(A) \approx 0.8953$.

Using this Lemma, we arrive at our main result.

THEOREM 6.4. *Under the assumptions of Lemma 2.1 and the restrictions on the partitioning of Ω_c detailed above, the classical Schwarz iteration (6.2) converges globally on an arbitrary number of subdomains.*

Proof. We will use Lemma 6.3 to show that $\rho(M_e) < 1$. The assumptions on the choice of the subdomains ensure the quantities (6.5) are non–negative, hence the matrix M_e is non–negative and, as mentioned previously, $\|M_e\|_\infty = 1$. The first, second last and last rows of M_e sum to a value less than one. Hence $\Psi_1(M_e)$ would zero rows and columns one, $I - 2$ and $I - 1$ of M_e. Then rows two, three, $I - 4$ and $I - 3$ of $\Psi_1(M_e)$ would have sums less than one, and hence $\Psi_1^2(M_e)$ would have zeros in rows and columns one to three and $I - 4$ to $I - 1$. It is easy to see that $\Psi_1^{\tilde{n}}(M_e) = 0$ for some $\tilde{n} < I - 1$. Therefore, by Lemma 6.3, $\rho(M_e) < 1$. A similar argument shows $\rho(M_{\hat{e}}) < 1$. □

This very general partitioning does not admit an explicit bound for the rate of convergence. An explicit contraction rate estimate is possible if we assume the overlaps are all of the same size. Analogous to Theorem 3.10 in [16] and Theorem 3.2 in Section 3, we assume that $r_i = s_i = r$ and $p_i = q_i = p$, and obtain the following explicit error estimate on I subdomains.

THEOREM 6.5. *The Schwarz iteration (6.2) on I subdomains with a common overlap ratio $r \in (0, 0.5]$ converges in the infinity norm and the iterates satisfy*

$$\max_{1 \le 2i \le I} \|x_{2i}^n(\xi) - x(\xi)\|_\infty \le \left(1 - 4r(1-r)\sin^2 \frac{\pi}{2(I+1)}\right)^n \frac{1}{\tilde{m}} \|e^0\|_2,$$

$$\max_{1 \le 2i+1 \le I} \|x_{2i+1}^n(\xi) - x(\xi)\|_\infty \le \left(1 - 4r(1-r)\sin^2 \frac{\pi}{2(I+1)}\right)^n \frac{1}{\tilde{m}} \|\hat{e}^0\|_2.$$

Proof. The proof is essentially given in [16] once we relate the point–wise error to the error measure (6.3) as in Theorem 3.2.

From the boundedness of M and (3.9), we have for each i and all ξ that $|x_{2i}^{2n+1}(\xi) - x(\xi)| \le \frac{1}{\tilde{m}}|e_{2i}^{2n+1}(\xi)|$. Now $|e_{2i}^{2n+1}|$ is linear and hence is bounded by the maximum of its boundary values. Therefore we have the sequence of inequalities:

$$|x_{2i}^{2n+1}(\xi) - x(\xi)| \le \frac{1}{\tilde{m}}|e_{2i}^{2n+1}(\xi)| \le \frac{1}{\tilde{m}} \max\left\{|e_{2i+1}^{2n}(\beta_{2i})|, |e_{2i-1}^{2n}(\alpha_{2i})|\right\}$$

$$\le \frac{1}{\tilde{m}}\|e^{2n}\|_\infty \le \frac{1}{\tilde{m}}\|e^{2n}\|_2 \le \frac{1}{\tilde{m}}\left(1 - 4r(1-r)\sin^2 \frac{\pi}{2(I+1)}\right)^n \|e^0\|_2.$$

The last inequality, giving a bound on the 2–norm of e^{2n} is given in [16]. Likewise we also obtain

$$|x_{2i+1}^{2n+1}(\xi) - x(\xi)| \le \frac{1}{\tilde{m}}|e_{2i+1}^{2n+1}(\xi)| \le \frac{1}{\tilde{m}} \max\left\{|e_{2i+2}^{2n}(\beta_{2i+1})|, |e_{2i}^{2n}(\alpha_{2i+1})|\right\}$$

$$\le \frac{1}{\tilde{m}}\|\hat{e}^{2n}\|_\infty \le \frac{1}{\tilde{m}}\|\hat{e}^{2n}\}_2 \le \frac{1}{\tilde{m}}\left(1 - 4r(1-r)\sin^2 \frac{\pi}{2(I+1)}\right)^n \|\hat{e}^0\|_2. \quad □$$

We see that the convergence rate deteriorates as the number of subdomains increases. The typical fix for this effect is the introduction of a coarse grid correction, like in the two level Schwarz methods, see for example [46].

7. Time Dependent Case. As mentioned in the Introduction, for a given time dependent function $u(t, x)$ a time dependent mesh transformation may be found by solving a nonlinear parabolic equation subject to appropriate initial and boundary conditions, see system (1.3). In this section, we consider the solution of (1.3) by first discretizing in time using an implicit method (backward Euler) and then solving the sequence of elliptic problems using a domain decomposition approach in the computational coordinate $\xi \in \Omega_c$.

Given the solution $x^{k-1}(\xi)$ at t_{k-1}, the backward Euler solution at time step k satisfies

$$x^k - \frac{\Delta t}{\tau}(M(x^k)x^k_\xi)_\xi = x^{k-1}, \quad x^k(0, t_k) = 0, \quad x^k(1, t_k) = 1. \tag{7.1}$$

For each time step $(k = 1, 2, \ldots)$ we solve (7.1) by the following classical, parallel, Schwarz iteration on two subdomains $\Omega_1 = (0, \beta)$ and $\Omega_2 = (\alpha, 1)$ with $\alpha < \beta$. For $n = 1, 2, \ldots$

$$x^{k,n}_1 - \frac{\Delta t}{\tau}(M(x^{k,n}_1)x^{k,n}_{1,\xi})_\xi = x^{k-1}, \text{ in } \Omega_1, \quad x^{k,n}_2 - \frac{\Delta t}{\tau}(M(x^{k,n}_2)x^{k,n}_{2,\xi})_\xi = x^{k-1}, \text{ in } \Omega_2,$$

$$x^{k,n}_1(0, t_k) = 0, \qquad\qquad\qquad x^{k,n}_2(\alpha, t_k) = x^{k,n-1}_1(\alpha, t_k),$$

$$x^{k,n}_1(\beta, t_k) = x^{k,n-1}_2(\beta, t_k), \qquad\qquad x^{k,n}_2(1, t_k) = 1. \tag{7.2}$$

The essential ingredient to show convergence of (7.2) is the classical maximum principle. A contraction rate is obtained by the method of supersolutions and the comparison principle [26] which we quote in the following Lemma.

LEMMA 7.1. *Suppose $Lu = au'' + bu' + cu$ is a linear, elliptic operator with $c \leq 0$ in a bounded domain Ω. Suppose that in Ω, $Lu \geq 0 (\leq 0)$ with $u \in C^2(\Omega) \cup C^0(\bar{\Omega})$. Then $\sup_\Omega u \leq \sup_{\partial\Omega} \max(u, 0)$ and $(\inf_\Omega u \geq \inf_{\partial\Omega} \min(u, 0))$.*

THEOREM 7.2. *Under the assumptions of Lemma 2.1, the iteration (7.2) converges for any time step $\Delta t > 0$ and for any mesh relaxation parameter $\tau > 0$. The convergence factor at the interfaces can be bounded by*

$$\rho_{time} = \frac{\sinh(\sqrt{\theta}\alpha)}{\sinh(\sqrt{\theta}\beta)} \frac{\sinh(\sqrt{\theta}(1-\beta))}{\sinh(\sqrt{\theta}(1-\alpha))} < 1, \quad \theta = \frac{\tau}{\Delta t}\frac{1}{\hat{m}}. \tag{7.3}$$

Proof. The separation of variables technique used in the proof of Theorem 3.2 is not possible. Instead we we proceed as in Section 6 and define an error measure

$$e^{k,n}_{1,2}(\xi) = \int_{x^{k,n}_{1,2}(\xi)}^{x^k(\xi)} M \, d\tilde{x}.$$

With this definition, we obtain the derivative

$$\frac{de^{k,n}_{1,2}}{d\xi} = M(x^k)\frac{dx^k}{d\xi} - M(x^{k,n}_{1,2})\frac{dx^{k,n}_{1,2}}{d\xi}, \tag{7.4}$$

and the mean value theorem for integrals implies

$$e^{k,n}_{1,2} = M(x^*_{1,2})(x^k - x^{k,n}_{1,2}), \tag{7.5}$$

for some $x^*_{1,2}$ between x^k and $x^{k,n}_{1,2}$.

Subtracting the equation for $x^{k,n}_1$ from the equation for x^k we obtain

$$x^k - x^{k,n}_1 - \frac{\Delta t}{\tau}\left(M(x^k)x^k_\xi - M(x^{k,n}_1)x^{k,n}_{1,\xi}\right)_\xi = 0,$$

17

and using the relations (7.4) and (7.5), the error functions are seen to satisfy

$$\frac{d^2 e_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}e_1^{k,n} = 0, \quad \text{in } \Omega_1, \qquad \frac{d^2 e_2^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_2^*)}e_2^{k,n} = 0, \quad \text{in } \Omega_2,$$

$$e_1^{k,n}(0, t_k) = 0, \qquad\qquad\qquad e_2^{k,n}(\alpha, t_k)) = e_1^{k,n-1}(\alpha, t_k),$$

$$e_1^{k,n}(\beta, t_k) = e_2^{k,n-1}(\beta, t_k), \qquad\qquad e_2^{k,n}(1, t_k) = 0.$$

Since $M, \tau, \Delta t$ are strictly positive, the error equations satisfy a maximum principle [26], and the required contraction results.

We obtain a contraction rate for the iteration (7.2) by construction of a supersolution for the errors on the subdomains. Assume $\tilde{e}_1^{k,n}$ solves the BVP

$$\frac{d^2 \tilde{e}_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{\hat{m}}\tilde{e}_1^{k,n} = 0, \quad \tilde{e}_1^{k,n}(0) = 0, \ \tilde{e}_1^{k,n}(\beta) = |e_2^{k,n-1}(\beta)|,$$

where we recall $\hat{m}$ denotes the upper bound on the nonlinear function M. The function $\tilde{e}_1^{k,n}$ may be found explicitly as

$$\tilde{e}_1^{k,n}(\xi) = |e_2^{k,n-1}(\beta)|\frac{\sinh(\sqrt{\theta}\xi)}{\sinh(\sqrt{\theta}\beta)}, \qquad \text{where} \qquad \theta = \frac{\tau}{\Delta t}\frac{1}{\hat{m}}.$$

We now show $\tilde{e}_1^{k,n}$ is a supersolution for $e_1^{k,n}$. To this end we define $d_1^{k,n} = e_1^{k,n} - \tilde{e}_1^{k,n}$, and routine manipulation shows $d_1^{k,n}$ satisfies

$$\frac{d^2 d_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}e_1^{k,n} + \frac{\tau}{\Delta t}\frac{1}{\hat{m}}\tilde{e}_1^{k,n} = 0, \ d_1^{k,n}(0) = 0, \ d_1^{k,n}(\beta) = e_2^{k,n-1}(\beta) - |e_2^{k,n-1}(\beta)|.$$
$$(7.6)$$

Adding and subtracting $\frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}\tilde{e}_1^{k,n}$ from (7.6) we see the $d_1^{k,n}$ satisfies the equation

$$\frac{d^2 d_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}d_1^{k,n} = \frac{\tau}{\Delta t}\left(\frac{1}{M(x_1^*)} - \frac{1}{\hat{m}}\right)\tilde{e}_1^{k,n}.$$

Since $\hat{m}$ is an upper bound for $M(x)$ the right hand side of the differential equation is greater than or equal to zero. The boundary value at $\xi = \beta$ is less than or equal to zero and the coefficient of $\tilde{e}_1^{k,n}$ in the differential equation is negative. Hence, Lemma 7.1 shows $d_1^{k,n} \leq 0$ or $e_1^{k,n} \leq \tilde{e}_1^{k,n}$ for all $\xi \in [0, \beta]$.

Now, the quantity $\tilde{d}_1^{k,n} = e_1^{k,n} + \tilde{e}_1^{k,n}$ satisfies

$$\frac{d^2 \tilde{d}_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}e_1^{k,n} - \frac{\tau}{\Delta t}\frac{1}{\hat{m}}\tilde{e}_1^{k,n} = 0, \ \tilde{d}_1^{k,n}(0) = 0, \ \tilde{d}_1^{k,n}(\beta) = e_2^{k,n-1}(\beta) + |e_2^{k,n-1}(\beta)|.$$

Again we add and subtract $\frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}\tilde{e}_1^{k,n}$ to show that $\tilde{d}_1^{k,n}$ satisfies the equation

$$\frac{d^2 \tilde{d}_1^{k,n}}{d\xi^2} - \frac{\tau}{\Delta t}\frac{1}{M(x_1^*)}\tilde{d}_1^{k,n} = \frac{\tau}{\Delta t}\left(\frac{1}{\hat{m}} - \frac{1}{M(x_1^*)}\right)\tilde{e}_1^{k,n}. \qquad (7.7)$$

The right hand side of the differential equation (7.7) is less than or equal to zero and the boundary conditions are greater than or equal to zero. Furthermore, the
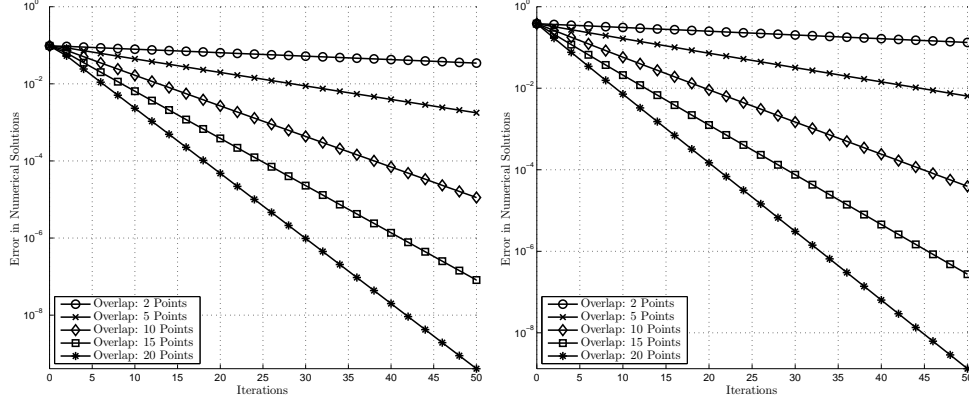
18

FIG. 8.1. *Convergence of classical Schwarz with varying overlap for $M(x) = x^2 + 1$ on the left, and for $M(x) = \sqrt{1 + u_x^2}$ with $u(x) = (1 - exp(20x))/(1 - exp(20))$ on the right.*

coefficient of $\tilde{d}_1^{k,n}$ is negative. The classical comparison principle then guarantees that $\tilde{d}_1^{k,n}(\xi) \geq 0$ or $e_1^{k,n}(\xi) \geq -\tilde{e}_1^{k,n}(\xi)$ for all $\xi \in [0, \beta]$. Hence, we have shown

$$|e_1^{k,n}(\xi)| \leq \tilde{e}_1^{k,n}(\xi) = |e_2^{k,n-1}(\beta)| \frac{\sinh(\sqrt{\theta}\xi)}{\sinh(\sqrt{\theta}\beta)}, \quad \text{for} \quad \xi \in \Omega_1.$$

Likewise, it possible to show

$$|e_2^{k,n}(\xi)| \leq |e_1^{k,n-1}(\alpha)| \frac{\sinh(\sqrt{\theta}(1 - \xi))}{\sinh(\sqrt{\theta}(1 - \alpha))}, \quad \text{for} \quad \xi \in \Omega_2.$$

Combining these relations we have

$$|e_1^{k,n+1}(\alpha)| \leq |e_1^{k,n-1}(\alpha)| \frac{\sinh(\sqrt{\theta}\alpha)}{\sinh(\sqrt{\theta}\beta)} \frac{\sinh(\sqrt{\theta}(1 - \beta))}{\sinh(\sqrt{\theta}(1 - \alpha))}.$$

If $\alpha < \beta$ then the contraction rate is bounded as shown in (7.3). $\square$

From the contraction estimate (7.3), we see that the algorithm slows down as the overlap becomes small, i.e. when $\alpha \to \beta$, we have $\rho \to 1$. We also see that the convergence of the DD iteration will improve as $\Delta t \to 0$, $\hat{m} \to 0$, or $\tau \to \infty$, since $\lim_{\theta \to \infty} \rho = 0$. Finally, the steady contraction rate is obtained in the opposite limit, since $\lim_{\theta \to 0} \rho = \frac{\alpha}{\beta} \frac{1-\beta}{1-\alpha}$.

8. Numerical Results. As a first example, we consider the solution of (1.4) with $M(x) = x^2 + 1$ using the classical parallel Schwarz algorithm (3.1) with two subdomains. In Figure 8.1 on the left, we illustrate the convergence history on subdomain one (in the maximum norm) versus the Schwarz iteration number – for varying amounts of overlap. The error is the difference between the discrete single domain solution and the discrete subdomain solutions. We see the convergence rate improve as the overlap increases. A more realistic example of grid generation is provided in Figure 8.1 on the right. We solve (1.4) with $M(x)$ chosen as the arclength monitor function $M(x) = \sqrt{1 + u_x^2}$ where

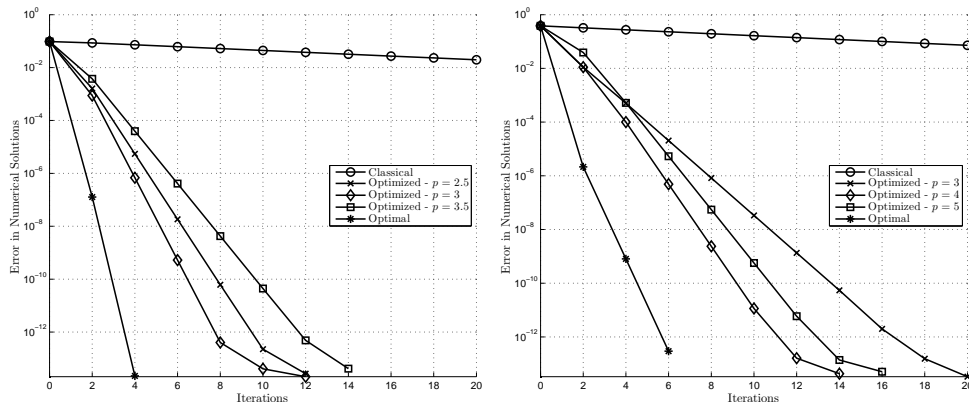$$u(x) = \frac{1 - e^{20x}}{1 - e^{20}}. \tag{8.1}$$

19

FIG. 8.2. *Comparison of classical, optimized and optimal Schwarz algorithms with $M(x) = x^2 + 1$ on the left, and $M(x) = \sqrt{1 + u_x^2}$ with $u(x) = (1 - exp(20x))/(1 - exp(20))$ on the right.*

This function is difficult to represent on a uniform grid in the physical coordinate x due to a boundary layer at $x = 1$.

In Figure 8.2 on the left, we compare the performance of the classical, optimized and optimal Schwarz algorithms applied to (1.4) with $M(x) = x^2 + 1$. We plot the error at every second iteration; this is consistent with the analysis of parallel Schwarz iterations which gives a contraction after two iterations on two subdomains. An exact antiderivative of M is used in the optimal transmission conditions (4.2–4.3). Clearly we see the benefit of the optimized and optimal Schwarz algorithms as compared to classical Schwarz. The results for optimized Schwarz are sensitive to the value of p – which may be tuned for performance. Inspecting (5.7) for $\alpha = 1/2$ we see that the contraction factor is minimized by choosing $p = 2\hat{m}$. Since $0 \leq x(\xi) \leq 1$, an estimate for the upper bound on M is $\hat{m} = 2$ which would give an optimal choice $p = 4$. If we were to start relatively close to the limit point $x(1/2) \approx 0.596$ then a good choice of p would be $p = 2(0.596^2 + 1) = 2.71$. These estimates agree, at least qualitatively, with the results shown. The optimal Schwarz algorithm is seen to be superior. Note, the optimal transmission conditions do not give convergence in two iterations for the discretized version of the algorithm. However, the error for the optimal Schwarz algorithm after two iterations does go to machine precision as the grid is refined. We repeat these experiments for the grid generation problem for $u(x)$ from (8.1) in Figure 8.2 on the right and see similar behavior. In this instance, numerical quadrature is used to approximate the optimal transmission conditions.

Theoretically, we have concentrated on the convergence of various Schwarz iterations to generate equidistributing grids. In practice, however, we are typically interested in the *quality* of the grids generated. In Table 8.1 we present the interpolation errors obtained when representing the function (8.1) on grids computed using nonlinear parallel, linearized parallel, optimal and optimized Schwarz algorithms. The number in brackets indicates the cumulative number of linear solves required by the end of that Schwarz iteration. The arclength monitor function was used to generate the grids. All iterations were started with an initial uniform grid. Although the parallel classical (nonlinear) Schwarz algorithm is slow to converge, we see that just a few iterations are needed to produce quality meshes, reducing the interpolation error by nearly a factor of 100 as compared to a uniform grid of the same size. The linearized

Iterations	0	2	4	6	8	10	∞
Nonlinear	0.3625	0.0520(5)	0.0498(10)	0.0478(15)	0.0462(21)	0.0448(27)	0.0366
Linearized	0.3625	0.1291(3)	0.1006(5)	0.0571(7)	0.0479(9)	0.0471(11)	0.0366
Optimized	0.3625	0.0416(9)	0.0367(23)	0.0366(30)	0.0366(36)	0.0366(40)	0.0366
Optimal	0.3625	0.0367(12)	0.0366(19)	0.0366(24)	0.0366(27)	0.0366(29)	0.0366

TABLE 8.1

Interpolation errors for the grids obtained by the various parallel Schwarz iterations. The numbers in brackets indicate the cumulative number of linear solves required.
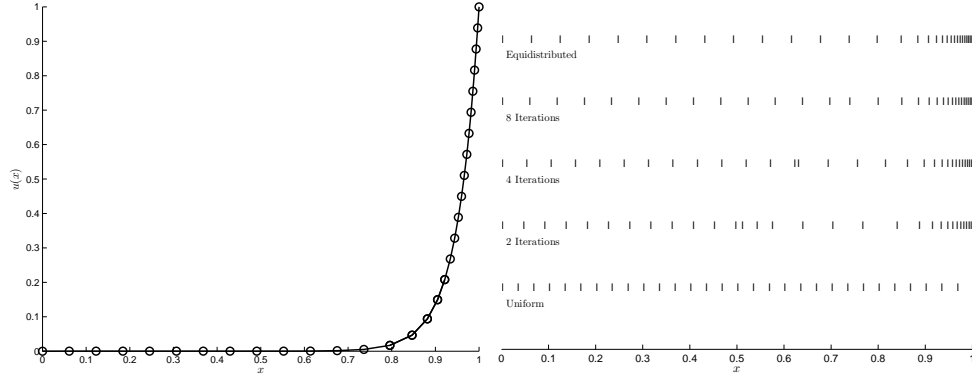

FIG. 8.3. *The function $u(x)$ from (8.1) on the equidistributing mesh (left) and the meshes obtained during the first 8 classical Schwarz iterations.*

Schwarz iteration takes many more DD iterations to reduce the interpolation error to the same level but is quite competitive if an appropriate work unit (linear solves for example) is used for comparison. The optimized and optimal algorithms achieve the best interpolation error possible in many fewer DD iterations as compared to the classical Schwarz algorithm.

Figure 8.3 illustrates the meshes obtained during the first few parallel Schwarz iterations. Only a few iterations are required to give a qualitatively correct grid.

In Figure 8.4 on the left, we compare four classical Schwarz variants: the parallel nonlinear and linearized Schwarz algorithms from Section 3 and the alternating versions from [22]. We choose the nonlinear function $M(x) = x^2 + 1$. In terms of the number of DD iterations, the alternating algorithms converge more quickly than the parallel counterparts. However, the alternating algorithms forgo easy parallelization. We also see that the nonlinear algorithms take fewer iterations to converge than the corresponding linearized version. In Figure 8.4 on the right, we again compare the four classical Schwarz variants mentioned above, but in this case we use $M(x)$ corresponding to the mesh generation problem for $u(x)$ given in (8.1). Here we display the convergence histories as a function of the number of required linear solves. Again we see the improved convergence of the alternating algorithms. In this case we see the utility of the linearized variants – outperforming the nonlinear algorithms.

The analysis of the linearized parallel Schwarz algorithm suggests a dependence on the shape of M, and in particular the ratio $\hat{m}/\check{m}$. To test this, in Figure 8.5 on the left, we experiment with $M(x) = C(x - 1/2)^2 + 1$. In this case, $\hat{m}/\check{m} \to \infty$ as $C \to \infty$. Indeed, we see that the performance of the linearized parallel Schwarz
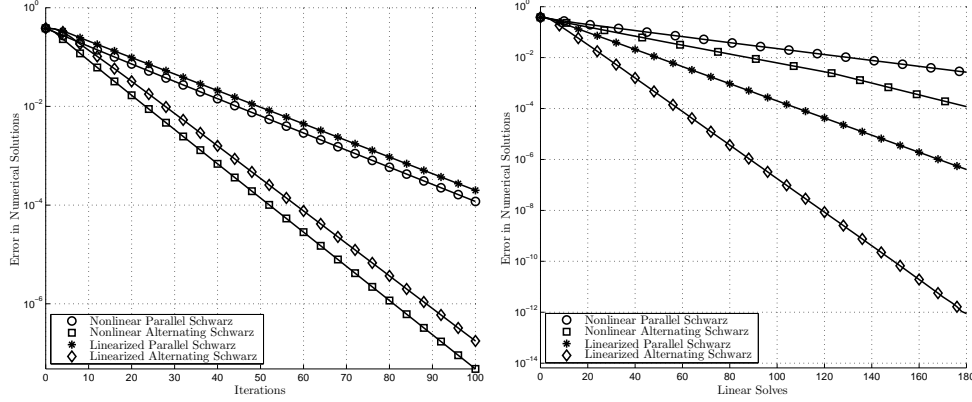
FIG. 8.4. *Comparison of nonlinear parallel, nonlinear alternating, linearized parallel and linearized alternating Schwarz algorithms for $M(x) = x^2 + 1$ on the left as a function of iteration numbers, and on the right for $M(x) = \sqrt{1 + u_x^2}$ with $u(x) = (1 - exp(20x))/(1 - exp(20))$, as a function of linear solves.*
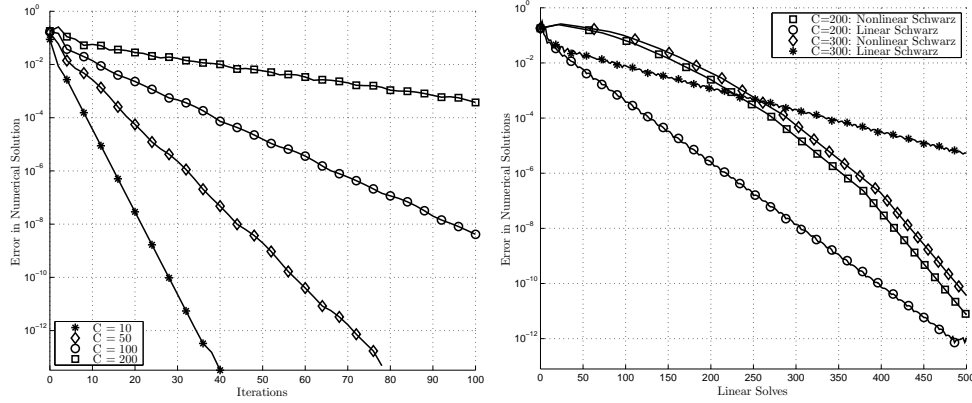

FIG. 8.5. *Linearized Parallel Schwarz for $M(x) = C(x - 1/2)^2 + 1$ for varying C on the left, and Nonlinear and Linearized Schwarz for $M(x) = C(x - 1/2)^2 + 1$ on the right.*

algorithm degrades as $C \to \infty$. In Figure 8.5 on the right, we use the same function M to see how changes in C affect both the linearized and nonlinear Schwarz algorithms. The results indicate that the linearized Schwarz algorithms are affected more as the problem becomes more difficult. Indeed, as the problem becomes more challenging ($C = 300$ for example) the nonlinear algorithm outperforms the linearized variant for small error tolerances.

To illustrate the DD approach for the generation of time dependent meshes we consider for $0 \leq x \leq 1$, $0 \leq t \leq 0.1$ the function

$$u(x, t) = \frac{1}{2} \left[1 - \tanh(c(t)(x - t - 0.4)) \right], \quad c(t) = 1 + \frac{499}{2} \left[1 + \tanh(10(t - 0.2)) \right].$$

This function has regions of rapid transition in space and time. A hyperbolic tangent profile develops and moves from left to right, as viewed in the physical coordinate x. We solve (MMPDE5), with M chosen as the arclength monitor function, using the Schwarz iteration described in Section 7. In Figure 8.6, we illustrate the subdomains
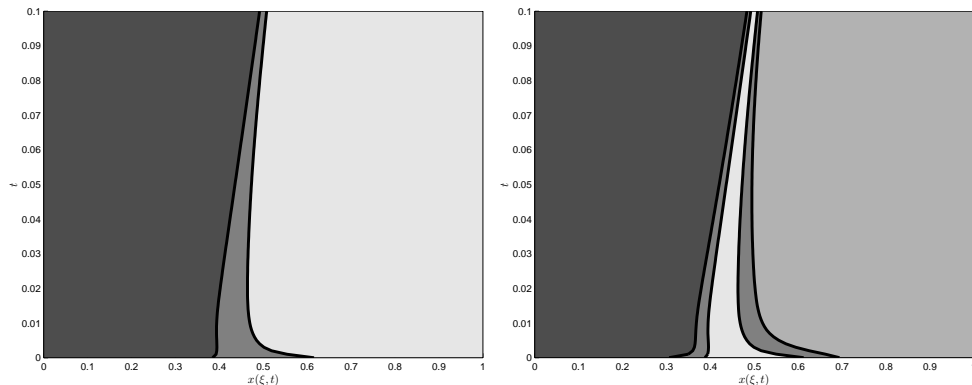
FIG. 8.6. *Two moving subdomains on the left, and three moving subdomains on the right.*

obtained in the physical coordinate using two (left) and three (right) overlapping subdomains. The medium grey in each figure denotes the overlap region. The subdomains and overlap are fixed (in both position and width) in the computational coordinate ξ. This gives rises to *moving* (time dependent) subdomains in the physical coordinate.

9. Conclusions. This paper provides a theoretical analysis of domain decomposition approaches for the parallel generation of equidistributing grids. Classical (nonlinear and linearized), optimal and optimized Schwarz algorithms were presented and studied. In the case of classical Schwarz the analysis was presented on an arbitrary, finite number of subdomains. Numerical results were presented which confirmed the theoretical findings. It is worthwhile to note that quality meshes, as determined by the resulting interpolation error, maybe obtained with a small number of DD iterations; convergence of the DD iteration may not be necessary. This is an interesting experimental result which requires more investigation in a practical setting where the mesh generation is coupled with the solution of a physical PDE.

Future work includes the extension of the optimal and optimized Schwarz algorithms for steady mesh generation to an arbitrary number of subdomains. The development of optimal and optimized algorithms for the generation of time dependent meshes is necessary. The development of a coarse grid correction to facilitate the parallel generation of equidistributing grids on a large number of cores or processors will be the subject of future investigations. The nonlinearity present in the grid generation problem presents some interesting challenges for the analysis of the discrete algorithms. This work is ongoing and will be presented elsewhere. Current work includes the analysis and implementation of the analogous DD approaches for equidistributing grid generation in two and three spatial dimensions. Although numerically it is relatively easy to extend to higher dimensions, the prescription of equidistributing grids in 2D/3D is not simply a trivial extension of the 1D BVP studied here. Hence a simple extension of the DD analysis presented in this paper will not be possible and further work is necessary.

REFERENCES

[1] Uri M. Ascher. DAEs that should not be solved. In *Dynamics of algorithms (Minneapolis, MN, 1997)*, volume 118 of *IMA Vol. Math. Appl.*, pages 55–67. Springer, New York, 2000.

[2] I. Babuška and W. C. Rheinboldt. Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.*, 15(4):736–754, 1978.

[3] Robert G. Bartle and Donald R. Sherbert. *Introduction to real analysis*. John Wiley & Sons Inc., New York, second edition, 1992.

[4] G. Beckett, J. A. Mackenzie, A. Ramage, and D. M. Sloan. Computational solution of two-dimensional unsteady PDEs using moving mesh methods. *J. Comput. Phys.*, 182(2):478–495, 2002.

[5] 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.

[6] Morton Bjorhus. *On Domain Decomposition, Subdomain Iteration, and Waveform Relaxation*. PhD thesis, Norwegian Institute of Technology, The University of Trondheim, 1995.

[7] 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.

[8] C. J. Budd, W. Huang, and R. D. Russell. Adaptivity with moving grids. *Acta Numer.*, 18:111–241, 2009.

[9] Xiao-Chuan Cai. Additive Schwarz algorithms for parabolic convection-diffusion equations. *Numer. Math.*, 60(1):41–61, 1991.

[10] 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).

[11] 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.

[12] 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.

[13] Leszek Demkowicz. *Computing with hp-adaptive finite elements. Vol. 1*. Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2007. One and two dimensional elliptic and Maxwell problems.

[14] Yana Di, Ruo Li, Tao Tang, and Pingwen Zhang. Moving mesh finite element methods for the incompressible Navier-Stokes equations. *SIAM J. Sci. Comput.*, 26(3):1036–1056, 2005.

[15] M. Dryja and W. Hackbusch. On the nonlinear domain decomposition method. *BIT*, 37(2):296–311, 1997.

[16] M. J. Gander and A. M. Stuart. Space–time continuous analysis of waveform relaxation for the heat equation. *SIAM J. Sci. Comput.*, 19(6):2014–2031, 1998.

[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. of 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 Journal of Numerical Analysis*, 41(5):1643–1681, 2003.

[22] Martin J. Gander and Ronald D. Haynes. Alternating and linearized alternating schwarz methods for equidistributing grids, 2012. Preprint, Submitted.

[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] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[27] Ronald D. Haynes. Recent advances in Schwarz waveform relaxation moving mesh methods – a new moving subdomain method. In *Domain decomposition methods in science and engineering XIX*, volume 78 of *Lect. Notes Comput. Sci. Eng.*, pages 253–260. Springer, Berlin, 2011.

[28] Ronald D. Haynes, Weizhang Huang, and Robert D. Russell. A moving mesh method for time-

dependent problems based on Schwarz waveform relaxation. In *Domain decomposition methods in science and engineering XVII*, volume 60 of *Lect. Notes Comput. Sci. Eng.*, pages 229–236. Springer, Berlin, 2008.

[29] Ronald D. Haynes and Robert D. Russell. A Schwarz waveform moving mesh method. *SIAM J. Sci. Comput.*, 29(2):656–673, 2007.

[30] W. Huang. Practical aspects of formulation and solution of moving mesh partial differential equations. *J. Comput. Phys.*, 171(2):753–775, 2001.

[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. Analysis of moving mesh partial differential equations with spatial smoothing. *SIAM J. Numer. Anal.*, 34(3):1106–1126, 1997.

[33] Weizhang Huang and Robert D. Russell. *Adaptive Moving Mesh Methods*, volume 174 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 2011.

[34] R. B. Kellogg. A nonlinear alternating direction method. *Math. Comp.*, 23:23–27, 1969.

[35] 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.

[36] S. H. Lui. On linear monotone iteration and Schwarz methods for nonlinear elliptic PDEs. *Numer. Math.*, 93(1):109–129, 2002.

[37] Shiu-Hong Lui. On monotone and Schwarz alternating methods for nonlinear elliptic PDEs. *M2AN Math. Model. Numer. Anal.*, 35(1):1–15, 2001.

[38] J. A. Mackenzie and W. R. Mekwi. An analysis of stability and convergence of a finite-difference discretization of a model parabolic PDE in 1D using a moving mesh. *IMA J. Numer. Anal.*, 27(3):507–528, 2007.

[39] J. A. Mackenzie and A. Nicola. A discontinuous Galerkin moving mesh method for Hamilton-Jacobi equations. *SIAM J. Sci. Comput.*, 29(6):2258–2282 (electronic), 2007.

[40] 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.

[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] D. W. Peaceman and H. H. Rachford, Jr. The numerical solution of parabolic and elliptic differential equations. *J. Soc. Indust. Appl. Math.*, 3:28–41, 1955.

[43] Linda R. Petzold. A description of DASSL: a differential/algebraic system solver. In *Scientific computing (Montreal, Que., 1982)*, IMACS Trans. Sci. Comput., I, pages 65–68. IMACS, New Brunswick, NJ, 1983.

[44] 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.

[45] Tao Tang. Moving mesh methods for computational fluid dynamics. In *Recent advances in adaptive computation*, volume 383 of *Contemp. Math.*, pages 141–173. Amer. Math. Soc., Providence, RI, 2005.

[46] Andrea Toselli and Olof Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2005.

[47] A. van Dam and P. A. Zegeling. A robust moving mesh finite volume method applied to 1D hyperbolic conservation laws from magnetohydrodynamics. *J. Comput. Phys.*, 216(2):526–546, 2006.

[48] A. van Dam and P. A. Zegeling. Balanced monitoring of flow phenomena in moving mesh methods. *Commun. Comput. Phys.*, 7(1):138–170, 2010.

[49] Baodong Zheng and Liancheng Wang. Spectral radius and infinity norm of matrices. *J. Math. Anal. Appl.*, 346(1):243–250, 2008.