DIRICHLET-NEUMANN AND NEUMANN-NEUMANN WAVEFORM RELAXATION ALGORITHMS FOR PARABOLIC PROBLEMS

MARTIN J. GANDER*, FELIX KWOK†, AND BANKIM C. MANDAL‡

Abstract. We present and analyze waveform relaxation variants of the Dirichlet-Neumann and Neumann-Neumann methods for parabolic problems. These methods are based on a non-overlapping spatial domain decomposition, and each iteration involves subdomain solves with Dirichlet boundary conditions followed by subdomain solves with Neumann boundary conditions. However, unlike for elliptic problems, each subdomain solve now involves a solution in space and time, and the interface conditions are also time-dependent. We show for the heat equation that when we consider finite time intervals, the Dirichlet-Neumann and Neumann-Neumann methods converge superlinearly for an optimal choice of the relaxation parameter, similar to the case of Schwarz waveform relaxation algorithms. Our analysis is based on Laplace transforms and detailed kernel estimates. The convergence rate depends on the size of the subdomains as well as the length of the time window. For any other choice of the relaxation parameter, convergence is only linear. We illustrate our results with numerical experiments.

1. Introduction.

Space-time parallel methods for the numerical solution of partial differential equations (PDE) is currently an active research topic, for a review, see [14]. This is driven by the increasing demand of high resolution simulation of complex systems, as well as the increasing availability of computing clusters with thousands of cores or more. One attractive way of speeding up the computation is to parallelize the solution process using domain decomposition (DD) methods. In such methods, one divides the computational domain into several subdomains and decouples the problem by making an initial guess on the data along subdomain interfaces; one then solves the subdomain problems in parallel, checks for discrepancies (e.g., non-smoothness) along subdomain interfaces, and iterates the process until a smooth solution is obtained. For an introduction to DD methods and their convergence properties, the reader may refer to the survey paper [9], as well as the volumes [43, 46] and the references therein. To parallelize the solution of time-dependent PDEs, the classical approach consists of discretizing in time to obtain a sequence of steady problems, which are in turn solved by DD methods; this is sometimes known as Rothe's method, in honor of the German analyst Erich Rothe. One drawback of this approach is that one is obliged to take the same time step across the whole domain, which can be very restrictive when the problem contains variable coefficients or multiple time scales.

A different approach consists of the so-called Waveform Relaxation (WR) methods: here, one solves the time-dependent problem approximately by regarding certain functions, e.g. the source terms, as fixed and integrating the remaining ODE or PDE. One then updates the frozen terms using the new approximate solution and iterates to obtain a consistent solution. WR methods have their origin in the 19th century, with the invention of Picard–Lindelöf iterations for proving the existence of local solutions to ODEs, see [42, 31]. Lelarasmee, Ruehli and Sangiovanni-Vincentelli [30] were the first to introduce WR as a parallel method for solving systems of ODEs: there, they decouple a feedback circuit into many subcircuits; the interactions between subcircuits are lagged by one iteration, thereby allowing the subcircuits to be simulated independently. The extension of WR methods to time-dependent PDEs was started independently in [19, 21]: in these papers, the authors use the DD paradigm to decompose the PDEs into several subdomain problems, each posed in space and time. At each iteration, one solves the space-time subproblem over the entire time horizon, before communicating interface data across subdomains. This approach has some advantages over Rothe's method, namely the more efficient bulk communication of interface data, which is now done

*Department of Mathematics, University of Geneva, Switzerland (martin.gander@unige.ch).
†Department of Mathematics, Hong Kong Baptist University, Hong Kong (felix_kwok@hkbu.edu.hk).
‡Department of Mathematical Sciences, Michigan Technological University, USA (bmandal@mtu.edu).

over the whole time horizon instead of per time step. It is not immediately obvious that the WR approach allows across-time parallelism in the same way that parareal [32] and parabolic multigrid [26] do. However, as Ong et al. [41] pointed out, it is possible to implement WR in a way that allows several iterations to run simultaneously on different parts of the time horizon, and without changing the mathematical properties of the algorithm. Thus, WR exposes opportunities for parallelism across time, and can be regarded as a space-time parallel method. We will explain this in more detail in Section 7.4, where we also show that the WR approach can lead to a net reduction in wall-clock time over Rothe's method, especially when the number of available cores is larger than can be exploited by spatial parallelism alone. This is in addition to the added flexibility of allowing for different spatial and temporal grids for each subdomain, which is only possible for WR methods; we show one such example in Section 7.3 and refer to [22] for a further discussion of these issues.

For WR methods applied to ODEs, we have two classical convergence results: (i) linear convergence on unbounded time intervals under some dissipation assumptions on the splitting ([38, 39, 27, 37]); and (ii) superlinear convergence for general, possibly nonlinear systems on bounded time intervals, assuming a Lipschitz condition on the splitting function ([38, 39, 1, 3]). For parabolic PDEs, Gander and Stuart [19] showed linear convergence of the overlapping Schwarz WR iteration for the heat equation on unbounded time intervals, with a rate depending on the size of the overlap. Giladi and Keller [21] proved superlinear convergence of the overlapping Schwarz WR method on bounded time intervals for the convection-diffusion equation. As is the case for stationary elliptic problems, the Schwarz WR method with classical (Dirichlet) interface conditions converges relatively slowly, except when the time window size is very short. A remedy is to use optimized interface conditions, which lead to much faster algorithms, see [16, 2] for parabolic problems, and [17, 15] for hyperbolic problems. This is in analogy with elliptic problems, where the use of optimized interface conditions also leads to large improvements, see for example [13] and references therein. In fact, it is straightforward to generalize any DD method formulated for steady problems into one for time-dependent problems using the WR approach, at least formally. A different class of space-time methods based on a discontinuous Galerkin discretization using space-time elements has been developed in the finite element community, see [44, 45].

Recently, we have been interested in studying the WR variants of two substructuring methods: the Dirichlet–Neumann (DN) algorithm, first considered by Bjørstad and Widlund [4] and further studied in [6, 36, 35]; and the Neumann–Neumann (NN) method, introduced by Bourgat et al. [5], see also [11, 29]. Our study is motivated by several reasons: in the elliptic case, such methods are closely related to FETI methods [12], which are some of the most commonly used and thoroughly tested DD methods in existence. Moreover, when used with a coarse grid, NN and FETI methods for elliptic problems scale very well in the mesh size and in the number of subdomains: it is shown in [29] that the NN-preconditioned system matrix has a condition number proportional to $H^{-2}(1+\log(H/h))^2$ without coarse grid and to $(1+\log(H/h))^2$ in case of a coarse grid correction, where h is the mesh size and H is the size of the subdomains. A systematic formulation of WR variants of the DN and NN methods does however not yet appear in the literature (two exceptions are [23] for a variant for control problems, and [25] for diffusion problems in mixed formulations). Thus, the main goals of this paper are as follows:

- To introduce the Dirichlet–Neumann Waveform Relaxation (DNWR) and Neumann–Neumann Waveform Relaxation (NNWR) methods for parabolic initial value problems, for a general decomposition into subdomains;
- To analyze the convergence of these methods for model problems in 1D, and in 2D for simple geometries;

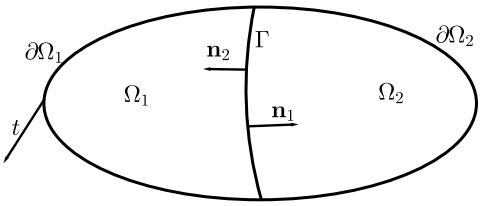

FIG. 2.1. *Splitting into two non-overlapping subdomains*

- To illustrate numerically the theoretical convergence rates, and to test if similar convergence behavior can also be observed for examples with spatially varying coefficients, and more general geometric decompositions, where our analysis does not apply.

Our analysis shows that the convergence speed of these methods depends on the number of subdomains, except if the time interval under consideration is shortened appropriately when the number of subdomains increases. Otherwise, a coarse grid correction is needed for true scalability, just like in the elliptic case. However, even without coarse grid correction, there is to our knowledge no convergence analysis of the DN and NN methods for initial value problems in the literature. This is because parabolic initial value problems are highly non-normal in the time direction, and existing techniques for estimating the condition number based on abstract Schwarz theory [46] cannot be used to analyze such methods, although they have been successful in dealing with elliptic problems in a very general setting.

In the remainder of this paper, we will formulate and analyze the new algorithms for the following parabolic equation on an open and bounded domain $\Omega \subset \mathbb{R}^d$, $0 < t < T$, $d = 1, 2, 3$:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \nabla \cdot (\kappa(\boldsymbol{x},t)\nabla u) + f(\boldsymbol{x},t), && \boldsymbol{x} \in \Omega,\ 0 < t < T, \\ u(\boldsymbol{x},0) &= u_0(\boldsymbol{x}), && \boldsymbol{x} \in \Omega, \\ u(\boldsymbol{x},t) &= g(\boldsymbol{x},t), && \boldsymbol{x} \in \partial\Omega,\ 0 < t < T, \end{aligned} \tag{1.1}$$

where $\kappa(\boldsymbol{x},t) \geq \kappa > 0$. In Section 2, we introduce the non-overlapping DNWR algorithm with two subdomains for the model problem (1.1), and we present convergence estimates for DNWR obtained for the special case of the one dimensional heat equation, $\kappa(\boldsymbol{x},t) = 1$, $d = 1$. In Section 3 we present the NNWR algorithm for multiple subdomains for the general problem (1.1), and present convergence estimates again for the one dimensional heat equation. Our convergence analysis shows that both the DNWR and NNWR algorithms converge superlinearly on finite time intervals, $T < \infty$. It is based on detailed, technical kernel estimates, which we show in Section 4. Section 5 contains the proofs of our main convergence results for both DNWR and NNWR. We then show in Section 6 how the analysis of the NNWR can be generalized to two spatial dimensions, and prove that the convergence estimates do not change for a particular decomposition into rectangular strips. We finally show numerical results in Section 7, which illustrate our analysis. We also test the algorithms in a few configurations not covered by our analysis.

2. The Dirichlet-Neumann Waveform Relaxation algorithm. To define the Dirichlet-Neumann WR algorithm for the model problem (1.1) on the space-time domain $\Omega \times (0,T)$ with Dirichlet data given on $\partial\Omega$, we assume that the spatial domain Ω is partitioned into two non-overlapping subdomains Ω_1 and Ω_2, as illustrated in Figure 2.1. We denote by u_i

the restriction of the solution u of (1.1) to Ω_i, $i = 1, 2$, and by $\boldsymbol{n}_i$ the unit outward normal for Ω_i on the interface $\Gamma := \partial\Omega_1 \cap \partial\Omega_2$. The Dirichlet-Neumann Waveform Relaxation algorithm consists of the following steps: given an initial guess $w^{(0)}(\boldsymbol{x}, t)$ along the interface $\Gamma \times (0, T)$, compute for $k = 1, 2, \ldots$ with $u_1^{(k)} = g$ on $\partial\Omega_1 \setminus \Gamma$ and $u_2^{(k)} = g$ on $\partial\Omega_2 \setminus \Gamma$ the approximations

(2.1)

$$\begin{aligned}
\partial_t u_1^{(k)} - \nabla \cdot \left(\kappa(\boldsymbol{x}, t)\nabla u_1^{(k)}\right) &= f, && \text{in } \Omega_1, \\
u_1^{(k)}(\boldsymbol{x}, 0) &= u_0(\boldsymbol{x}), && \text{in } \Omega_1, \\
u_1^{(k)} &= w^{(k-1)}, && \text{on } \Gamma,
\end{aligned}
\qquad
\begin{aligned}
\partial_t u_2^{(k)} - \nabla \cdot \left(\kappa(\boldsymbol{x}, t)\nabla u_2^{(k)}\right) &= f, && \text{in } \Omega_2, \\
u_2^{(k)}(\boldsymbol{x}, 0) &= u_0(\boldsymbol{x}), && \text{in } \Omega_2, \\
\partial_{\boldsymbol{n}_2} u_2^{(k)} &= -\partial_{\boldsymbol{n}_1} u_1^{(k)}, && \text{on } \Gamma,
\end{aligned}$$

and then update the value along the interface using

$$w^{(k)}(\boldsymbol{x}, t) = \theta u_2^{(k)}\big|_{\Gamma\times(0,T)} + (1 - \theta)w^{(k-1)}(\boldsymbol{x}, t), \tag{2.2}$$

$\theta \in (0, 1]$ being a relaxation parameter. Now the main goal of the analysis is to study how the error $w^{(k-1)}(\boldsymbol{x}, t) - u\big|_{\Gamma\times(0,T)}$ converges to zero, and by linearity it suffices to consider the so called error equations, with $f(\boldsymbol{x}, t) = 0$, $g(\boldsymbol{x}, t) = 0$, $u_0(\boldsymbol{x}) = 0$ in (2.1), and examine how $w^{(k-1)}(\boldsymbol{x}, t)$ converges to zero as $k \to \infty$.

We now present convergence estimates for algorithm (2.1)-(2.2) for the special case of the heat equation, $\kappa(\boldsymbol{x}, t) = 1$, on the one dimensional domain $\Omega = (-a, b)$ with subdomains $\Omega_1 = (-a, 0)$ and $\Omega_2 = (0, b)$. Our convergence analysis is based on Laplace transforms. We define the Laplace transform of a function $u(x, t)$ with respect to time t as

$$\hat{u}(x, s) := \mathcal{L}\{u(x, t)\} := \int_0^\infty e^{-st}u(x, t)\, dt, \tag{2.3}$$

where s is a complex variable. If $\mathcal{L}\{u(x, t)\} = \hat{u}(x, s)$, then the inverse Laplace transform of $\hat{u}(x, s)$ is defined by

$$\mathcal{L}^{-1}\{\hat{u}(x, s)\} := u(x, t), \qquad t \geq 0, \tag{2.4}$$

and maps the Laplace transform of a function back to the original function. For more information on Laplace transforms, see [10, 40].

After a Laplace transform, the DNWR algorithm (2.1)-(2.2) for the error equations in the one dimensional heat equation setting becomes

(2.5)

$$\begin{aligned}
(s - \partial_{xx})\hat{u}_1^{(k)} &= 0 && \text{on } (-a, 0), \\
\hat{u}_1^{(k)}(-a, s) &= 0, \\
\hat{u}_1^{(k)}(0, s) &= \hat{w}^{(k-1)}(s),
\end{aligned}
\qquad
\begin{aligned}
(s - \partial_{xx})\hat{u}_2^{(k)} &= 0 && \text{on } (0, b), \\
\partial_x\hat{u}_2^{(k)}(0, s) &= \partial_x\hat{u}_1^{(k)}(0, s), \\
\hat{u}_2^{(k)}(b, s) &= 0,
\end{aligned}$$

followed by the updating step

$$\hat{w}^{(k)}(s) = \theta\hat{u}_2^{(k)}(0, s) + (1 - \theta)\hat{w}^{(k-1)}(s). \tag{2.6}$$

Solving the two-point boundary value problems in the Dirichlet and Neumann step in (2.5), we get

$$\hat{u}_1^{(k)}(x, s) = \frac{\hat{w}^{(k-1)}(s)}{\sinh(a\sqrt{s})} \sinh\left((x + a)\sqrt{s}\right), \tag{2.7}$$

$$\hat{u}_2^{(k)}(x, s) = \hat{w}^{(k-1)}(s)\frac{\coth(a\sqrt{s})}{\cosh(b\sqrt{s})} \sinh((x - b)\sqrt{s}). \tag{2.8}$$

By induction, we therefore find for the updating step the relation

$$\hat{w}^{(k)}(s) = \left(1 - \theta - \theta \tanh(b\sqrt{s}) \coth(a\sqrt{s})\right)^k \hat{w}^{(0)}(s), \quad k = 1, 2, 3, \ldots \tag{2.9}$$

THEOREM 2.1 (Convergence of DNWR for $a = b$). *When the subdomains are of the same size, $a = b$ in (2.5)-(2.6), the DNWR algorithm converges linearly for $0 < \theta < 1$, $\theta \neq 1/2$. For $\theta = 1/2$, it converges in two iterations. Convergence is independent of the time window size T.*

Proof. For $a = b$, equation (2.9) reduces to $\hat{w}^{(k)}(s) = (1 - 2\theta)^k \hat{w}^{(0)}(s)$, which has the simple back transform $w^{(k)}(t) = (1 - 2\theta)^k w^{(0)}(t)$. Thus the convergence is linear for $0 < \theta < 1$, $\theta \neq 1/2$. If $\theta = 1/2$, we have $w^{(1)}(t) = 0$, and hence one more iteration produces the desired solution on the entire domain. ∎

Having treated the simple case where the subdomains are of the same size, $a = b$, we focus now on the more interesting case where $a \neq b$. Defining

$$F(s) := \tanh(b\sqrt{s}) \coth(a\sqrt{s}) - 1 = \frac{\sinh((b-a)\sqrt{s})}{\sinh(a\sqrt{s}) \cosh(b\sqrt{s})}, \tag{2.10}$$

the recurrence relation (2.9) can be rewritten as

$$\hat{w}^{(k)}(s) = \begin{cases} \left((1 - 2\theta) - \theta F(s)\right)^k \hat{w}^{(0)}(s), & \theta \neq 1/2, \\ (-1)^k 2^{-k} F^k(s) \hat{w}^{(0)}(s), & \theta = 1/2. \end{cases} \tag{2.11}$$

Note that for $\mathrm{Re}(s) > 0$, $F(s)$ is $\mathcal{O}(s^{-p})$ for every positive p, which can be seen as follows: setting $s = re^{i\vartheta}$, we obtain for $a \geq b$ the bound

$$|s^p F(s)| \leq \left| \frac{s^p}{\cosh(b\sqrt{s})} \right| \leq \frac{2r^p}{\left| e^{b\sqrt{r/2}} - e^{-b\sqrt{r/2}} \right|} \to 0 \quad \text{as } r \to \infty,$$

and for $a < b$, we get the bound

$$|s^p F(s)| \leq \left| \frac{s^p}{\sinh(a\sqrt{s})} \right| \leq \frac{2r^p}{\left| e^{a\sqrt{r/2}} - e^{-a\sqrt{r/2}} \right|} \to 0 \quad \text{as } r \to \infty.$$

Therefore, by [10, p. 178], $F(s)$ is the Laplace transform of an infinitely differentiable function $f_1(t)$, which is the reason why we introduced $F(s)$ in (2.10). We now define

$$f_k(t) := \mathcal{L}^{-1}\left\{F^k(s)\right\} \quad \text{for } k = 1, 2, \ldots. \tag{2.12}$$

In what follows, we study the special case $\theta = 1/2$, when $w^{(k)}$ is given by a convolution of $w^{(0)}$ with the analytic function f_k (see (4.1) for the definition of convolution). For θ not equal to $1/2$, different techniques are required to analyze the behavior of the DNWR algorithm, and this will be done in a future paper. We also have to consider two cases: $a > b$, which means that the Dirchlet subdomain is bigger than Neumann subdomain, and $a < b$, when the Neumann subdomain is bigger than the Dirichlet subdomain. We have the following two convergence results, whose proofs will be given in Section 5.

THEOREM 2.2 (Convergence of DNWR for $a > b$). *If $\theta = 1/2$ and the Dirichlet subdomain is larger than the Neumann subdomain, then the error of the DNWR algorithm (2.5)-(2.6) satisfies for $t \in (0, \infty)$ the linear convergence estimate*

$$\|w^{(k)}\|_{L^\infty(0,\infty)} \leq \left(\frac{a-b}{2a}\right)^k \|w^0\|_{L^\infty(0,\infty)}. \tag{2.13}$$

On a finite time interval $t \in (0,T)$, the DNWR method converges superlinearly with the estimate

$$\|w^{(k)}\|_{L^\infty(0,T)} \leq \left(\frac{a-b}{a}\right)^k \operatorname{erfc}\left(\frac{kb}{2\sqrt{T}}\right) \|w^{(0)}\|_{L^\infty(0,T)}. \tag{2.14}$$

THEOREM 2.3 (Convergence of DNWR for $a < b$). *If $\theta = 1/2$ and the Dirichlet subdomain is smaller than the Neumann subdomain, then the error of the DNWR algorithm (2.5)-(2.6) over $t \in (0,\infty)$ satisfies the linear bound*

$$\|w^{(2k)}\|_{L^\infty(0,\infty)} \leq \left(\frac{b-a}{2a}\right)^{2k} \|w^{(0)}\|_{L^\infty(0,\infty)}. \tag{2.15}$$

For a finite time interval $t \in (0,T)$, the DNWR converges superlinearly with the estimate

$$\|w^{(2k)}\|_{L^\infty(0,T)} \leq \left(\frac{\sqrt{2}}{1-e^{-\frac{(2k+1)a^2}{T}}}\right)^{2k} e^{-k^2a^2/T} \|w^{(0)}\|_{L^\infty(0,T)}. \tag{2.16}$$

REMARK 2.1. *The linear estimate (2.15) does not always imply convergence, because $b-a$ can be larger than $2a$. In other words, when $b > 3a$, i.e., when the Neumann subdomain is much larger than the Dirichlet one, it is not clear, at least from the expression of the estimate (2.15), whether the iteration converges to zero as $k \to \infty$. In that case, as a remedy, one should switch the interface conditions and solve a Dirichlet problem on the larger subdomain.*

REMARK 2.2. *Note that the factor multiplying $e^{-k^2a^2/T}$ in the estimate (2.16) is an increasing function of k in general, since $\frac{\sqrt{2}}{1-e^{-(2k+1)a^2/T}} > 1$. Thus, the bound (2.16) may increase initially for small iteration numbers k, before the factor $e^{-k^2a^2/T}$ starts dominating and causing the bound to decrease to zero superlinearly, see the right panel of Figure 3.2. To estimate the turning point, let us fix an integer $l > 0$ and consider the behavior of the algorithm for iteration numbers $k > 2l$. Then by writing $\sigma = T/a^2$ and $\alpha = e^{-l/\sigma}$, we can bound the convergence factor by*

$$\begin{aligned}\left(\frac{\sqrt{2}}{1-e^{-\frac{(2k+1)a^2}{T}}}\right)^{2k} e^{-k^2a^2/T} &\leq \left(\frac{\sqrt{2}}{1-e^{-\frac{2k+1}{\sigma}}}\right)^{2k} e^{-2kl/\sigma} e^{-k(k-2l)/\sigma} \\ &\leq \left(\frac{\sqrt{2}e^{-l/\sigma}}{1-e^{-\frac{2k}{\sigma}}}\right)^{2k} e^{-k(k-2l)/\sigma} \leq \underbrace{\left(\frac{\sqrt{2}\alpha}{1-\alpha^4}\right)^{2k}}_{=(*)} e^{-\frac{(k-2l)^2}{\sigma}}.\end{aligned}$$

Thus, if $\sqrt{2}\alpha/(1-\alpha^4) < 1$, then the factor $(*)$ is less than 1 and the bound contracts superlinearly for $k > 2l$. This is true whenever $\alpha < \alpha_0$, where $\alpha_0 \approx 0.6095$ is the unique positive root of $\psi(\alpha) = \alpha^4 + \sqrt{2}\alpha - 1$. Hence, we get superlinear convergence for $k > 2l > 0.99T/a^2$.

3. The Neumann-Neumann Waveform Relaxation algorithm. We now introduce the NNWR algorithm for the model problem (1.1) for multiple subdomains; for the case of two subdomains in 1D, see [28]. For the well-posedness of this method, we refer to [24]. Suppose Ω is partitioned into non-overlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$, as illustrated in Figure 3.1. For $i = 1,\ldots,N$ set $\Gamma_i := \partial\Omega_i \setminus \partial\Omega$, $\Lambda_i := \{j \in \{1,\ldots,N\} :$

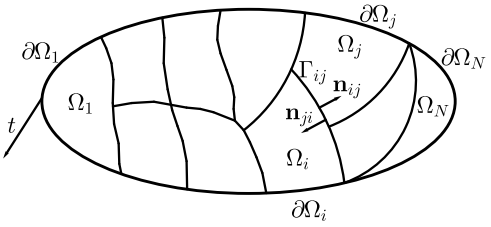

FIG. 3.1. *Splitting into many non-overlapping subdomains*

$\Gamma_i \cap \Gamma_j$ has nonzero measure$\}$ and $\Gamma_{ij} := \partial\Omega_i \cap \partial\Omega_j$, so that the interface of Ω_i can be rewritten as $\Gamma_i = \bigcup_{j\in\Lambda_i} \Gamma_{ij}$. We denote by $\boldsymbol{n}_{ij}$ the unit outward normal for Ω_i on the interface Γ_{ij}.

The NNWR algorithm starts with an initial guess $w_{ij}^{(0)}(\boldsymbol{x},t)$ along the interfaces $\Gamma_{ij} \times (0,T)$, $j \in \Lambda_i$, $i = 1,\ldots,N$, and then performs the following two-step iteration: at each iteration k, one first solves Dirichlet problems on each Ω_i in parallel,

$$
\begin{aligned}
\partial_t u_i^{(k)} - \nabla\cdot\left(\kappa(\boldsymbol{x},t)\nabla u_i^{(k)}\right) &= f, && \text{in } \Omega_i,\\
u_i^{(k)}(\boldsymbol{x},0) &= u_0(\boldsymbol{x}), && \text{in } \Omega_i,\\
u_i^{(k)} &= g, && \text{on } \partial\Omega_i \setminus \Gamma_i,\\
u_i^{(k)} &= w_{ij}^{(k-1)}, && \text{on } \Gamma_{ij}, j \in \Lambda_i.
\end{aligned}
\tag{3.1}
$$

One then solves Neumann problems on all subdomains in parallel,

$$
\begin{aligned}
\partial_t \psi_i^{(k)} - \nabla\cdot\left(\kappa(\boldsymbol{x},t)\nabla \psi_i^{(k)}\right) &= 0, && \text{in } \Omega_i,\\
\psi_i^{(k)}(\boldsymbol{x},0) &= 0, && \text{in } \Omega_i,\\
\psi_i^{(k)} &= 0, && \text{on } \partial\Omega_i \setminus \Gamma_i,\\
\partial_{\boldsymbol{n}_{ij}} \psi_i^{(k)} &= \partial_{\boldsymbol{n}_{ij}} u_i^{(k)} + \partial_{\boldsymbol{n}_{ji}} u_j^{(k)}, && \text{on } \Gamma_{ij}, j \in \Lambda_i.
\end{aligned}
\tag{3.2}
$$

The interface values are then updated with the formula

$$
w_{ij}^{(k)}(\boldsymbol{x},t) = w_{ij}^{(k-1)}(\boldsymbol{x},t) - \theta\left(\psi_i^{(k)}\big|_{\Gamma_{ij}\times(0,T)} + \psi_j^{(k)}\big|_{\Gamma_{ij}\times(0,T)}\right), \tag{3.3}
$$

where $\theta \in (0,1]$ is a relaxation parameter*.

As in the case of the DNWR algorithm, we prove our results first for the one dimensional heat equation on the domain $\Omega := (0,L)$ with boundary conditions $u(0,t) = g_0(t)$ and $u(L,t) = g_L(t)$; for a special case in two spatial dimensions, see Section 6. We decompose Ω into non-overlapping subdomains $\Omega_i := (x_{i-1}, x_i)$, $i = 1,\ldots,N$, and define the subdomain length $h_i := x_i - x_{i-1}$, and $h_{\min} := \min_{1\le i\le N} h_i$. Our initial guess is denoted by $\left\{w_i^{(0)}(t)\right\}_{i=1}^{N-1}$ on the interfaces x_i. By linearity, we again study the error equations,

*The second step of NNWR computes a correction ψ for the jump in the normal derivative of u, so θ is indeed a relaxation parameter: it is like when solving $A\mathbf{x} = \mathbf{b}$ with the matrix splitting $A = M - N$, $M\tilde{\mathbf{x}} = N\mathbf{x}_n + \mathbf{b}$ followed by the relaxation $\mathbf{x}_{n+1} = (1-\theta)\mathbf{x}_n + \theta\tilde{\mathbf{x}}$, which can be rewritten as a correction, $\mathbf{x}_{n+1} = (1-\theta)\mathbf{x}_n + \theta(M^{-1}N\mathbf{x}_n + M^{-1}\mathbf{b}) = \mathbf{x}_n + \theta(M^{-1}N\mathbf{x}_n + M^{-1}\mathbf{b} - \mathbf{x}_n)$, which is now of the form (3.3).

TABLE 3.1
Comparison of estimates of various known WR methods.

Methods	2 subdomains	N equal subdomains
Classical Schwarz WR [20, 16]	$\operatorname{erfc}\left(\frac{k\delta}{\sqrt{T}}\right)$	$2^k\operatorname{erfc}\left(\frac{k\delta}{2\sqrt{T}}\right)$
OSWR with overlap δ [2, 16]	$1-C(\Delta x)^{\frac{1}{3}}$ if $\beta \geq \frac{4}{3}$ $1-C(\Delta x)^{\frac{\beta}{4}}$ if $\beta < \frac{4}{3}$	convergence proof only
OSWR without overlap [2, 16]	$1-C(\Delta t)^{\frac{1}{4}}$	convergence proof only
DNWR	$\left(\frac{a-b}{a}\right)^k \operatorname{erfc}\left(\frac{kb}{2\sqrt{T}}\right)$	$(N-2)^k \operatorname{erfc}\left(\frac{kh}{2\sqrt{T}}\right)$
NNWR	$\left(\frac{(a-b)^2}{ab}\right)^k \operatorname{erfc}\left(\frac{kb}{\sqrt{T}}\right)$	$\left(\frac{\sqrt{6}}{1-e^{-\frac{(2k+1)h^2}{T}}}\right)^{2k} e^{\frac{-k^2h^2}{T}}$

$f=0$, $g_0=g_L=0$ and $u_0=0$, which leads with $w_0^{(k)}(t)=w_N^{(k)}(t)=0$ for all k to

$$
\begin{aligned}
&\partial_t u_i^{(k)} - \partial_{xx} u_i^{(k)} = 0, && \text{in } \Omega_i, && \partial_t \psi_i^{(k)} - \partial_{xx}\psi_i^{(k)} = 0, && \text{in } \Omega_i,\\
&u_i^{(k)}(x,0) = 0, && \text{in } \Omega_i, && \psi_i^{(k)}(x,0) = 0, && \text{in } \Omega_i,\\
&u_i^{(k)}(x_{i-1},t) = w_{i-1}^{(k-1)}(t), && && -\partial_x\psi_i^{(k)}(x_{i-1},t) = (\partial_x u_{i-1}^{(k)} - \partial_x u_i^{(k)})(x_{i-1},t),\\
&u_i^{(k)}(x_i,t) = w_i^{(k-1)}(t), && && \partial_x\psi_i^{(k)}(x_i,t) = (\partial_x u_i^{(k)} - \partial_x u_{i+1}^{(k)})(x_i,t),
\end{aligned}
\tag{3.4}
$$

except for the first and last subdomains, where in the Neumann step the Neumann conditions are replaced by homogeneous Dirichlet conditions at the physical boundaries. The new interface values for the next step are then defined as

$$
w_i^{(k)}(t) = w_i^{(k-1)}(t) - \theta\left(\psi_i^{(k)}(x_i,t) + \psi_{i+1}^{(k)}(x_i,t)\right). \tag{3.5}
$$

We have the following convergence result for NNWR:

THEOREM 3.1 (Convergence of NNWR). *For $\theta = 1/4$ and $T > 0$ fixed, the NNWR algorithm (3.4)–(3.5) converges superlinearly with the estimate*

(3.6)

$$
\max_{1\leq i\leq N-1} \|w_i^{(k)}\|_{L^\infty(0,T)} \leq \left(\frac{\sqrt{6}}{1-e^{-\frac{(2k+1)h_{\min}^2}{T}}}\right)^{2k} e^{-k^2h_{\min}^2/T} \max_{1\leq i\leq N-1} \|w_i^{(0)}\|_{L^\infty(0,T)}.
$$

The proof of Theorem 3.1 will also be given in Section 5. Similar to the DNWR case, the parameter choice $\theta = 1/4$ leads to convergence in two iterations in the two-subdomain case when the subdomain sizes are equal, see [28] for a full discussion. Thus, it is reasonable to expect superlinear convergence when the subdomains have different sizes. The case $\theta \neq 1/4$ requires different analysis techniques and is the subject of further studies.

Before proceeding further with the analysis and proofs of the theorems, we give a summary of the known bounds in Table 3.1 for other WR algorithms to compare the effectiveness of the newly found DNWR and NNWR algorithms. In Figure 3.2 we compare the theoretical estimates of various known WR methods with that of the DNWR and NNWR methods from Table 3.1, where $\Delta t = (\Delta x)^\beta$. For the Schwarz WR methods, we use an overlap δ of length $4\Delta x$, where $\Delta x = 1/50$, $\beta \approx 1.4114$, subdomain lengths $a = 3, b = 2$ and $C = 2$. For the many subdomains case, we take $\delta = 12\Delta x, N = 5, h = 1$. We observe that the estimates of DNWR and NNWR algorithms indicate faster convergence than the other methods for this particular choice.

REMARK 3.1. *Table 3.1 states that NNWR converges in about half as many iterations as DNWR. However, in comparison to DNWR, the NNWR has to solve twice the number*

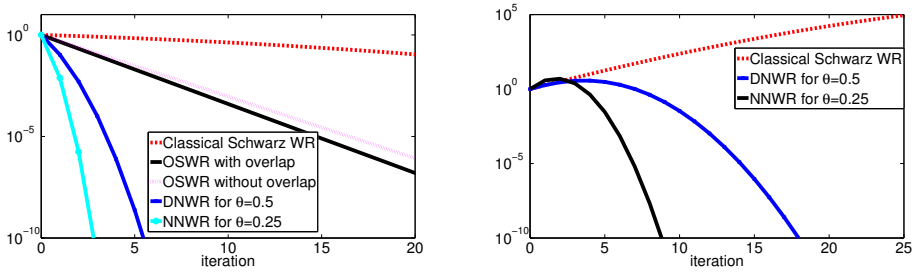

FIG. 3.2. *Comparison of estimates of various WR algorithms on the left for two subdomains, and on the right for many subdomains*

of subproblems (once for Dirichlet subproblems, and once for Neumann subproblems) on each subdomain at each iteration. Therefore, the computational cost is almost double for the NNWR compared to the DNWR algorithm at each step, so the actual amount of computation necessary for convergence is similar for both methods. Nonetheless, the NNWR has the advantage of being easily generalizable to multiple subdomains, unlike DNWR (but see [18]).

REMARK 3.2. *Just as we mentioned in Remark 2.2, the error bound for NNWR may increase for the first few iterations for large T, before superlinear convergence kicks in. This behavior is reminiscent of the behavior of $\|A^k\|$ for non-normal matrices A as k increases. This is not a coincidence: in fact, one can rewrite the discretized parabolic PDE as a large linear system whose unknowns are the solution values at all grid points and at all time steps. Because of causality, this is a block lower triangular matrix with the same diagonal blocks if the diffusion coefficients do not vary in time. NNWR can then be seen as a stationary iterative method applied to this non-normal, non-diagonalizable matrix. Thus, we observe the same hump behavior in the initial iterates as for powers of Jordan blocks.*

REMARK 3.3. *For the model problem (1.1) posed on the time interval $(0,T)$, one may consider applying NNWR to the whole interval $(0,T)$, or to a sequence of smaller time* windows *$(0,T/M)$, $(T/M,2T/M)$, $\ldots$, $((M-1)T/M,T)$. Since the total number of time steps across all time windows remains constant, the best time window size should be chosen to minimize the* running time per time step*; for a simple computational model of NNWR, this is given by*

$$R = K_w \left(T_s + T_t + \frac{T_m}{N_t} \right),$$

where N_t is the number of time steps in the time window, K_w is the number of WR iterations required for convergence, T_s is the solution time per time step *of a subdomain problem, T_t is the time required to transfer interface data over one time step (since the amount of data to be transferred is proportional to the number of time steps), and T_m is the set-up time for data transfer (once per iteration). The convergence estimate (3.6) can help us choose a time window size to minimize running time. Let $B(\alpha,k,t)$ be the contraction factor of NNWR, as defined in (4.5). Then for a moderately-sized positive constant c (i.e., c is $O(1)$), we have*

$$B(\alpha,ck,ct) \lesssim \left(\frac{2}{1-e^{-(k+1/c)\alpha^2/t}} \right)^{ck} e^{-ck^2\alpha^2/4t} \approx (B(\alpha,k,t))^c.$$

We can use this scaling property to choose our time window size adaptively as follows. Suppose we have previously chosen a time window T_{init} and obtained a convergence curve $E(k)$,

where k is the iteration count and $E(k)$ represents the error at step k. Then for a given tolerance ϵ and a given time window size T_{new}, the number of iterations K_w required for convergence is approximately $K_w \approx ck^*$, where $c = T_{new}/T_{init}$ and k^* satisfies

$$\frac{\log \epsilon}{\log(E(k^*))} = c.$$

Thus, for each T_{new}, there is a corresponding K_w, so one can solve the minimization problem by trying different values of $N_t = T_{new}/\Delta t$. This technique will allow us to adjust the time window size as the solver progresses, without needing access to the machine-dependent parameters T_s, T_t and T_m.

4. Kernel estimates. The convergence results given in Theorems 2.2, 2.3 and 3.1 are based on technical estimates of kernels arising in the Laplace transform of the DNWR and NNWR algorithms. We present in this section the precise estimates needed.

4.1. Properties of Laplace transforms. We start with several elementary properties of non-negative functions and their Laplace transforms. We define the convolution of two functions $g, w : (0,\infty) \to \mathbb{R}$ by

$$(g * w)(t) := \int_0^t g(t-\tau)w(\tau)d\tau. \tag{4.1}$$

LEMMA 4.1. *Let g and w be two real-valued functions in $(0,\infty)$ with $\hat{w}(s) = \mathcal{L}\{w(t)\}$ the Laplace transform of w. Then for $t \in (0,T)$, we have the following properties:*

1. *If $g(t) \geq 0$ and $w(t) \geq 0$, then $(g * w)(t) \geq 0$.*
2. *$\|g * w\|_{L^1(0,T)} \leq \|g\|_{L^1(0,T)}\|w\|_{L^1(0,T)}$.*
3. *$\big|(g * w)(t)\big| \leq \|g\|_{L^\infty(0,T)} \int_0^T \big|w(\tau)\big|d\tau$.*
4. *$\int_0^t w(\tau)d\tau = \mathcal{L}^{-1}\left(\frac{\hat{w}(s)}{s}\right)$.*

Proof. The proofs follow directly from the definitions. ∎

LEMMA 4.2. *Let $w(t)$ be a continuous and L^1-integrable function on $(0,\infty)$ with $w(t) \geq 0$ for all $t \geq 0$, and $\hat{w}(s) = \mathcal{L}\{w(t)\}$ be its Laplace transform. Then, for $\tau > 0$, we have the bound*

$$\int_0^\tau |w(t)|dt \leq \lim_{s\to 0+} \hat{w}(s). \tag{4.2}$$

Proof. With the definition of the Laplace transform (2.3) and using positivity, we have

$$\begin{aligned}\int_0^\tau |w(t)|dt = \int_0^\tau w(t)dt \leq \int_0^\infty w(t)dt = \int_0^\infty \lim_{s\to 0+} e^{-st}w(t)dt \\ = \lim_{s\to 0+}\int_0^\infty e^{-st}w(t)dt = \lim_{s\to 0+}\hat{w}(s),\end{aligned}$$

where the dominated convergence theorem was used to exchange the order of limit and integration. ∎

LEMMA 4.3. *Let $w(t)$ be a function whose Laplace transform $\hat{w}(s) = \int_0^\infty w(t)e^{-st}\,dt$ is defined for all $s \in D \subset \mathbb{C}$. Then the function*

$$v(t) = w(t)e^{-\eta t}$$

has a Laplace transform at s whenever $s + \eta \in D$, and $\mathcal{L}\{v\}(s) = \hat{w}(s+\eta)$.

Proof. We have

$$\mathcal{L}\{v\}(s) = \int_0^\infty w(t)e^{-\eta t}e^{-st}\,dt = \int_0^\infty w(t)e^{-(s+\eta)t}\,dt = \hat{w}(s+\eta),$$

whenever the latter is defined. ∎

4.2. Positivity. In order to use Lemma 4.2 in our analysis, we have to show positivity of the inverse transforms of kernels appearing in the DNWR and NNWR iteration. These results are established in the following lemma.

LEMMA 4.4. *Let $\beta > \alpha \geq 0$ and s be a complex variable. Then, for $t \in (0,\infty)$*

$$\varphi(t) := \mathcal{L}^{-1}\left\{\frac{\sinh(\alpha\sqrt{s})}{\sinh(\beta\sqrt{s})}\right\} \geq 0 \quad \textit{and} \quad \psi(t) := \mathcal{L}^{-1}\left\{\frac{\cosh(\alpha\sqrt{s})}{\cosh(\beta\sqrt{s})}\right\} \geq 0.$$

Proof. We first prove that φ and ψ are well-defined and continous functions on $(0,\infty)$. Setting $s = re^{i\vartheta}$, a short calculation shows that for $\beta > \alpha \geq 0$ and for every positive p

$$\left|\frac{s^p\sinh(\alpha\sqrt{s})}{\sinh(\beta\sqrt{s})}\right| \leq r^p \cdot \left|\frac{e^{\alpha\sqrt{r/2}} + e^{-\alpha\sqrt{r/2}}}{e^{\beta\sqrt{r/2}} - e^{-\beta\sqrt{r/2}}}\right| \to 0 \qquad \text{as } r \to \infty,$$

so by [10, p. 178], its inverse Laplace transform exists and is continuous (in fact, infinitely differentiable). Thus, φ is a continuous function. A similar argument holds for ψ.

Next, we prove the non-negativity of φ and ψ by noting that these kernels are related to solutions of the heat equation. Let us consider the heat equation $u_t - u_{xx} = 0$ on $(0,\beta)$ with initial condition $u(x,0) = 0$ and boundary conditions $u(0,t) = 0$, $u(\beta,t) = g(t)$. If g is non-negative, then by the maximum principle, this boundary value problem has a non-negative solution $u(\alpha,t)$ for all $\alpha \in [0,\beta]$, $t > 0$. Now performing a Laplace transform of the heat equation in time, we obtain the transformed solution along $x = \alpha$ to be

$$\hat{u}(\alpha,s) = \hat{g}(s)\frac{\sinh(\alpha\sqrt{s})}{\sinh(\beta\sqrt{s})} \quad \Longrightarrow \quad u(\alpha,t) = \int_0^t g(t-\tau)\varphi(\tau)d\tau.$$

We now prove that $\varphi(t) \geq 0$ by contradiction: suppose $\varphi(t_0) < 0$ for some $t_0 > 0$. Then by the continuity of φ, there exists $\delta > 0$ such that $\varphi(\tau) < 0$ for $\tau \in (t_0-\delta, t_0+\delta)$. Now for $t > t_0 + \delta$, we choose a non-negative g as follows:

$$g(\zeta) = \begin{cases} 1, & \zeta \in (t-t_0-\delta, t-t_0+\delta)\,, \\ 0, & \text{otherwise.} \end{cases}$$

Then $u(\alpha,t) = \int_{t_0-\delta}^{t_0+\delta} g(t-\tau)\varphi(\tau)d\tau = \int_{t_0-\delta}^{t_0+\delta}\varphi(\tau)d\tau < 0$, which is a contradiction, and hence φ must be non-negative. To prove the result for ψ, we use again the heat equation $u_t - u_{xx} = 0$, $u(x,0) = 0$, but on the domain $(-\beta,\beta)$ and with boundary conditions $u(-\beta,t) = u(\beta,t) = g(t)$. Using a Laplace transform in time gives as solution at $x = \alpha$

$$\hat{u}(\alpha,s) = \hat{g}(s)\frac{\cosh(\alpha\sqrt{s})}{\cosh(\beta\sqrt{s})},$$

and hence a similar argument as in the first case proves that ψ is also non-negative. ∎

4.3. Specific kernel estimates. The following lemma contains specific estimates for the inverse Laplace transform of two kernels in terms of infinite sums.

LEMMA 4.5. *For $k = 1, 2, 3, \ldots$, we have the identities*

$$\mathcal{L}^{-1}\left(\operatorname{cosech}^k(\alpha\sqrt{s})\right) = 2^k \sum_{m=0}^{\infty} \binom{m+k-1}{m} \frac{(2m+k)\alpha}{\sqrt{4\pi t^3}} e^{-(2m+k)^2\alpha^2/4t}, \tag{4.3}$$

$$\mathcal{L}^{-1}\left(\frac{\operatorname{cosech}^k(\alpha\sqrt{s})}{s}\right) = 2^k \sum_{m=0}^{\infty} \binom{m+k-1}{m} \operatorname{erfc}\left(\frac{(2m+k)\alpha}{2\sqrt{t}}\right). \tag{4.4}$$

In particular, both functions are positive for $t > 0$. Moreover, we have the estimate

$$B(\alpha, k, t) := \mathcal{L}^{-1}\left(\frac{\operatorname{cosech}^k(\alpha\sqrt{s})}{s}\right) \le \left(\frac{2}{1 - e^{-(k+1)\alpha^2/t}}\right)^k e^{-k^2\alpha^2/4t}. \tag{4.5}$$

Proof. First, we show that the Laplace transform of the right-hand side of (4.3) gives $\operatorname{cosech}^k(\alpha\sqrt{s})$. This is the same as showing that

$$\operatorname{cosech}^k(\alpha\sqrt{s}) = 2^k \int_0^\infty e^{-st} \sum_{m=0}^{\infty} \binom{m+k-1}{m} \frac{(2m+k)\alpha}{\sqrt{4\pi t^3}} e^{-(2m+k)^2\alpha^2/4t}\, dt. \tag{4.6}$$

We would like to integrate term by term, i.e., interchange the sum and the integral. This is possible as long as we have absolute summability, i.e., if

$$\sum_{m=0}^{\infty} \binom{m+k-1}{m} \int_0^\infty \left| e^{-st} \frac{(2m+k)\alpha}{\sqrt{4\pi t^3}} e^{-(2m+k)^2\alpha^2/4t} \right| dt < \infty.$$

We prove this for $\operatorname{Re}(s) \ge s_0 > 0$. From Oberhettinger [40], we know that

$$\mathcal{L}^{-1}\left(e^{-\lambda\sqrt{s}}\right) = \frac{\lambda}{\sqrt{4\pi t^3}} e^{-\lambda^2/4t}, \quad \lambda > 0. \tag{4.7}$$

So

$$\int_0^\infty \left| e^{-st} \frac{(2m+k)\alpha}{\sqrt{4\pi t^3}} e^{-(2m+k)^2\alpha^2/4t} \right| dt = e^{-(2m+k)\alpha\sqrt{\operatorname{Re}(s)}} \le e^{-(2m+k)\alpha\sqrt{s_0}}.$$

Thus, using the binomial series

$$\frac{1}{(1-z)^\beta} = \sum_{m \ge 0} \binom{m+\beta-1}{m} z^m \qquad \text{for } |z| < 1 \tag{4.8}$$

with $z = e^{-2\alpha\sqrt{s_0}} < 1$, we get

$$\begin{aligned} &\sum_{m=0}^{\infty} \binom{m+k-1}{m} \int_0^\infty \left| e^{-st} \frac{(2m+k)\alpha}{\sqrt{4\pi t^3}} e^{-(2m+k)^2\alpha^2/4t} \right| dt \\ &\qquad \le \frac{e^{-k\alpha\sqrt{s_0}}}{(1 - e^{-2\alpha\sqrt{s_0}})^k} = \frac{\operatorname{cosech}^k(\alpha\sqrt{s_0})}{2^k} < \infty. \end{aligned}$$

Therefore, term-by-term integration is possible, and the same calculation without the absolute value proves (4.6). Moreover, since each term in the right hand side of (4.3) is positive, the sum must be positive.

The second identity is established in a similar way, except we replaced (4.7) with

$$\mathcal{L}^{-1}\left(\frac{1}{s}e^{-\lambda\sqrt{s}}\right) = \operatorname{erfc}\left(\frac{\lambda}{2\sqrt{t}}\right), \quad \lambda > 0, \tag{4.9}$$

see Oberhettinger [40]. This function is again positive; to obtain a bound of the form (4.5) without summations, we first note the estimate

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt = \frac{2}{\sqrt{\pi}}\int_0^\infty e^{-(x+t)^2}dt \le \frac{2}{\sqrt{\pi}}e^{-x^2}\int_0^\infty e^{-t^2}dt \le e^{-x^2},$$

which is valid for all $x \ge 0$. Substituting this into the formula (4.4) leads to

$$\begin{aligned}
B(\alpha,k,t) &\le 2^k \sum_{m=0}^\infty \binom{m+k-1}{m} \exp\left(-\frac{(2m+k)^2\alpha^2}{4t}\right) \\
&= 2^k e^{-k^2\alpha^2/4t} \sum_{m=0}^\infty \binom{m+k-1}{m} \exp\left(-\frac{(m^2+km)\alpha^2}{t}\right) \\
&\le 2^k e^{-k^2\alpha^2/4t} \sum_{m=0}^\infty \binom{m+k-1}{m} \exp\left(-\frac{m(k+1)\alpha^2}{t}\right) \\
&\le \left(\frac{2}{1-e^{-(k+1)\alpha^2/t}}\right)^k e^{-k^2\alpha^2/4t},
\end{aligned}$$

where we used the binomial series (4.8) with $z = e^{-(k+1)\alpha^2/t}$ for the last inequality. □

5. Proofs of the main theorems. We now prove the main convergence results for the DNWR and NNWR algorithms stated in Section 2 and 3.

5.1. Proof of Theorem 2.2 for DNWR. The goal is to obtain L^∞ estimates for the DNWR error $w^{(k)}(t)$ in the case of $\theta = 1/2$ and $a > b$, i.e., when the Dirichlet domain is larger. From (2.11), we see that for $\theta = 1/2$, we have

$$\hat{w}^{(k)}(s) = (-1)^k 2^{-k} F^k(s)\hat{w}^{(0)}(s),$$

where $F(s) = \frac{\sinh((b-a)\sqrt{s})}{\sinh(a\sqrt{s})\cosh(b\sqrt{s})}$. Denoting the inverse transform of $F^k(s)$ by $f_k(t)$, cf. (2.12), we use part 3 of Lemma 4.1 to deduce that

$$\left|w^{(k)}(t)\right| = \left|2^{-k}\left(w^{(0)} * f_k\right)(t)\right| \le 2^{-k}\|w^{(0)}\|_{L^\infty(0,T)}\int_0^T |f_k(\tau)|\,d\tau. \tag{5.1}$$

The estimates (2.13) and (2.14) now follow from different ways of bounding $\int_0^T |f_k(\tau)|\,d\tau$. In particular, we will prove that

(i) $\displaystyle\int_0^T |f_k(\tau)|\,d\tau \le \left(\frac{a-b}{a}\right)^k$ for all $T > 0$. Since this bound is independent of T, this leads to the linear estimate (2.13)

$$\|w^{(k)}\|_{L^\infty(0,\infty)} \le \left(\frac{a-b}{2a}\right)^k \|w^{(0)}\|_{L^\infty(0,\infty)};$$

(ii) $\displaystyle\int_0^T |f_k(\tau)|\,d\tau \le 2^k\left(\frac{a-b}{a}\right)^k \operatorname{erfc}(kbT^{-1/2}/2)$ for all $T > 0$, which leads to the superlinear estimate (2.14)

$$\|w^{(k)}\|_{L^\infty(0,T)} \le \left(\frac{a-b}{a}\right)^k \operatorname{erfc}\left(\frac{kb}{2\sqrt{T}}\right)\|w^{(0)}\|_{L^\infty(0,T)}.$$

Proof of (i): We first consider the case $k = 1$, where we have

$$\mathcal{L}(-f_1(t)) = -F(s) = \frac{\sinh((a-b)\sqrt{s})}{\sinh(a\sqrt{s})} \cdot \frac{1}{\cosh(b\sqrt{s})}.$$

Since $a > b$, we see by Lemma 4.4 that both factors are Laplace transforms of non-negative functions. Thus, $-f_1(t)$ is the convolution of two non-negative functions and is therefore non-negative. Convolving $-f_1(t)$ with itself k times shows that $(-1)^k f_k(t)$ is also non-negative, so by Lemma 4.2, we have

$$\int_0^T |f_k(\tau)|\, d\tau = \int_0^T (-1)^k f_k(\tau)\, d\tau \le \lim_{s\to 0+} (-1)^k F^k(s) = \left(\frac{a-b}{a}\right)^k.$$

The linear estimate (2.13) then follows from this inequality.

Proof of (ii): Here we rewrite $f_k(t)$ as the convolution $(-1)^k(p_k * q_k)(t)$, where

$$\mathcal{L}(p_k(t)) = \hat{p}^k(s) = \frac{\sinh^k((a-b)\sqrt{s})}{\sinh^k(a\sqrt{s})} \quad \text{and} \quad \mathcal{L}(q_k(t)) = \hat{q}^k(s) = \frac{1}{\cosh^k(b\sqrt{s})}.$$

Using the same reasoning as in part (i), we obtain the bound

$$\|p_k(t)\|_{L^\infty(0,T)} \le \left(\frac{a-b}{a}\right)^k,$$

so part 3 of Lemma 4.1 implies

$$\int_0^T |f_k(\tau)|\, d\tau \le \left(\frac{a-b}{a}\right)^k \int_0^T |q_k(\tau)|\, d\tau. \tag{5.2}$$

Since $\int_0^T |q_k(\tau)|\, d\tau$ is difficult to calculate directly, we will use the following trick: we choose a non-negative function $r_k(t)$ such that (a) $r_k(t) \ge q_k(t)$ for all t, and (b) its antiderivative is known. This then allows us to bound $\int_0^T q_k(\tau)\, d\tau$ by the antiderivative of r_k evaluated at T. Let $r_k(t) = \mathcal{L}^{-1}(2^k e^{-kb\sqrt{s}})$. To show that $r_k(t) - q_k(t)$ is a non-negative function, we calculate its Laplace transform:

$$\begin{aligned}\mathcal{L}\left\{r_k(t) - q_k(t)\right\} &= 2^k e^{-kb\sqrt{s}} - \frac{2^k}{(e^{b\sqrt{s}} + e^{-b\sqrt{s}})^k} \\ &= \frac{2^k((1+e^{-2b\sqrt{s}})^k - 1)}{(e^{b\sqrt{s}} + e^{-b\sqrt{s}})^k} = \sum_{j=1}^k \binom{k}{j} e^{-2jb\sqrt{s}} \hat{q}^k(s).\end{aligned}$$

From (4.7), we know that $\mathcal{L}^{-1}(e^{-2jb\sqrt{s}})$ is a non-negative function, and so is $q_k(t)$. Thus, each term in the sum above corresponds to the convolution of two non-negative functions, which is non-negative. Thus, $r_k(t) - q_k(t) \ge 0$, so we deduce that

$$\int_0^T q_k(\tau)\, d\tau \le \int_0^T r_k(\tau)\, d\tau = \mathcal{L}^{-1}\left(\frac{2^k e^{-kb\sqrt{s}}}{s}\right) = 2^k \operatorname{erfc}\left(\frac{kb}{2\sqrt{T}}\right),$$

where we expressed the second integral as an inverse Laplace transform using Lemma 4.1, part 4, which we then evaluated using (4.9). Substituting into (5.2) leads to the required estimate.

5.2. Proof of Theorem 2.3 for DNWR. Recall that the recurrence in Laplace space reads

$$\hat{w}^{(2k)}(s) = 2^{-2k} F^{2k}(s) \hat{w}^{(0)}(s), \qquad \text{where} \quad F(s) = \frac{\sinh((b-a)\sqrt{s})}{\sinh(a\sqrt{s})\cosh(b\sqrt{s})}.$$

The linear estimate (2.15) has already been given in [34], so we will focus on the superlinear estimate. Here, we assume that $a < b$, i.e., the Dirichlet domain is smaller than the Neumann domain. Unlike in Theorem 2.2, it is possible to have $|b-a| > a$ in this case, so we can no longer argue that $\sinh((b-a)\sqrt{s})/\sinh(a\sqrt{s})$ is the Laplace transform of a regular function. (We say that this is an *improper pairing* between the numerator and denominator.) Here, we use a different argument: the trick is to write $F^k(s)$ as a product of two kernels, one of which corresponds to a superlinearly decaying term, and the other to a function whose L^1 norm grows at most geometrically. We write

$$F^{2k}(s) = \frac{1}{\sinh^{2k}(a\sqrt{s})} \cdot \left(\frac{\sinh^2((b-a)\sqrt{s})}{\cosh^2(b\sqrt{s})} \right)^k =: \Phi^{2k}(s) \cdot \Psi^k(s),$$

where

$$\Psi(s) = \frac{\sinh^2((b-a)\sqrt{s})}{\cosh^2(b\sqrt{s})} = \underbrace{\frac{\cosh^2((b-a)\sqrt{s})}{\cosh^2(b\sqrt{s})}}_{=:\mathcal{L}(r_1(t))} - \underbrace{\frac{1}{\cosh^2(b\sqrt{s})}}_{=:\mathcal{L}(r_2(t))}.$$

Observe that r_1 and r_2 are regular functions by Lemma 4.4, since $0 < b-a < b$. Denoting the inverse Laplace transform of $\Psi^k(s)$ by $\psi_k(t) := \mathcal{L}^{-1}(\Psi^k(s))$, we obtain by Lemma 4.2

$$\int_0^T |\psi_1(\tau)|\, d\tau \le \int_0^T |r_1(\tau)| + |r_2(\tau)|\, d\tau \le \lim_{s\to 0+} \frac{\cosh^2((b-a)\sqrt{s})}{\cosh^2(b\sqrt{s})} + \lim_{s\to 0+} \frac{1}{\cosh^2(b\sqrt{s})} = 2.$$

By part 2 of Lemma 4.1, we get

$$\|\psi_k\|_{L^1(0,T)} \le \|\psi_1\|^k_{L^1(0,T)} \le 2^k. \tag{5.3}$$

Next, denote the inverse Laplace transform of $\Phi^{2k}(s)$ by $\phi_{2k}(t) := \mathcal{L}^{-1}(\Phi^{2k}(s))$. Then Lemma 4.5 shows that ϕ_{2k} is a non-negative function, so part 4 of Lemma 4.1 implies

$$\|\phi_{2k}\|_{L^1(0,T)} = \mathcal{L}^{-1}\left(\frac{1}{s} \operatorname{cosech}^{2k}(a\sqrt{s}) \right) = B(a, 2k, T) \le \left(\frac{2}{1 - e^{-(2k+1)a^2/T}} \right)^{2k} e^{-k^2a^2/T}, \tag{5.4}$$

where $B(a,k,t)$ is defined in (4.5). Combining (5.3), (5.4) and the definition of $\hat{w}^{(2k)}(s)$, we find the required estimate

$$\|w^{(2k)}\|_{L^\infty(0,T)} \le \left(\frac{\sqrt{2}}{1 - e^{-(2k+1)a^2/T}} \right)^{2k} e^{-k^2a^2/T} \|w^{(0)}\|_{L^\infty(0,T)}.$$

5.3. Proof of Theorem 3.1 for NNWR. The proof of Theorem 3.1 will be divided into several parts. We first derive the recurrence relation between interface values in Laplace space.

LEMMA 5.1. *Let $w_i^{(k)}(t)$ be the values along the interfaces $\{x_i\}_{i=1}^{N-1}$ at the kth iteration of the NNWR algorithm (3.4)–(3.5), and let $\hat{w}_i^{(k)}(s)$ be its Laplace transform. Then we have the recurrence*

$$\hat{\mathbf{w}}^{(k)} = A(s)\hat{\mathbf{w}}^{(k-1)},$$

where $\hat{\mathbf{w}}^{(k)} = (\hat{w}_1^{(k)}, \ldots, \hat{w}_{N-1}^{(k)})^T$*, and* A *is a pentadiagonal matrix such that the following recurrences are satisfied:*

(5.5)
$$\hat{w}_1^{(k)} = -\frac{1}{4}\left(\hat{w}_1^{(k-1)}\left(\frac{\gamma_1\gamma_2 - \sigma_1\sigma_2}{\sigma_1\sigma_2} + \frac{\sigma_1\gamma_2 - \gamma_1\sigma_2}{\gamma_1\sigma_2}\right) + \frac{\hat{w}_2^{(k-1)}}{\sigma_2}\left(\frac{\gamma_3}{\sigma_3} - \frac{\sigma_1}{\gamma_1}\right) - \frac{\hat{w}_3^{(k-1)}}{\sigma_2\sigma_3}\right),$$

(5.6)
$$\hat{w}_i^{(k)} = -\frac{1}{4}\left(\hat{w}_i^{(k-1)}\left(\frac{2(\gamma_i\gamma_{i+1} - \sigma_i\sigma_{i+1})}{\sigma_i\sigma_{i+1}}\right) + \frac{\hat{w}_{i+1}^{(k-1)}}{\sigma_{i+1}}\left(\frac{\gamma_{i+2}}{\sigma_{i+2}} - \frac{\gamma_i}{\sigma_i}\right) + \frac{\hat{w}_{i-1}^{(k-1)}}{\sigma_i}\left(\frac{\gamma_{i-1}}{\sigma_{i-1}} - \frac{\gamma_{i+1}}{\sigma_{i+1}}\right) - \frac{\hat{w}_{i+2}^{(k-1)}}{\sigma_{i+1}\sigma_{i+2}} - \frac{\hat{w}_{i-2}^{(k-1)}}{\sigma_i\sigma_{i-1}}\right), \quad i = 2, \ldots, N-2,$$

(5.7)
$$\hat{w}_{N-1}^{(k)} = -\frac{1}{4}\left(\hat{w}_{N-1}^{(k-1)}\left(\frac{\gamma_{N-1}\gamma_N - \sigma_{N-1}\sigma_N}{\sigma_{N-1}\sigma_N} + \frac{\sigma_N\gamma_{N-1} - \gamma_N\sigma_{N-1}}{\gamma_N\sigma_{N-1}}\right) + \frac{\hat{w}_{N-2}^{(k-1)}}{\sigma_{N-1}}\left(\frac{\gamma_{N-2}}{\sigma_{N-2}} - \frac{\sigma_N}{\gamma_N}\right) - \frac{\hat{w}_{N-3}^{(k-1)}}{\sigma_{N-1}\sigma_{N-2}}\right).$$

In the above formulas, we used the notation $\sigma_i = \sinh(h_i\sqrt{s})$, $\gamma_i = \cosh(h_i\sqrt{s})$.

Proof. We start by applying the Laplace transform to the homogeneous Dirichlet subproblems in (3.4), and obtain

$$s\hat{u}_i - \hat{u}_{i,xx} = 0, \quad \hat{u}_i(x_{i-1}, s) = \hat{w}_{i-1}(s), \quad \hat{u}_i(x_i, s) = \hat{w}_i(s),$$

for $i = 2, \ldots, N-1$. These subdomain problems have the solutions

$$\hat{u}_i(x, s) = \frac{1}{\sinh(h_i\sqrt{s})}\left(\hat{w}_i(s)\sinh\left((x - x_{i-1})\sqrt{s}\right) + \hat{w}_{i-1}(s)\sinh\left((x_i - x)\sqrt{s}\right)\right).$$

Next we apply the Laplace transform to the Neumann subproblems (3.2) for subdomains not touching the physical boundary, and obtain

$$\hat{\psi}_i(x, s) = C_i(s)\cosh\left((x - x_{i-1})\sqrt{s}\right) + D_i(s)\cosh\left((x_i - x)\sqrt{s}\right),$$

where the notation $\sigma_i := \sinh(h_i\sqrt{s})$ and $\gamma_i := \cosh(h_i\sqrt{s})$ gives

$$C_i = \frac{1}{\sigma_i}\left(\hat{w}_i\left(\frac{\gamma_i}{\sigma_i} + \frac{\gamma_{i+1}}{\sigma_{i+1}}\right) - \frac{\hat{w}_{i-1}}{\sigma_i} - \frac{\hat{w}_{i+1}}{\sigma_{i+1}}\right),$$
$$D_i = \frac{1}{\sigma_i}\left(\hat{w}_{i-1}\left(\frac{\gamma_i}{\sigma_i} + \frac{\gamma_{i-1}}{\sigma_{i-1}}\right) - \frac{\hat{w}_{i-2}}{\sigma_{i-1}} - \frac{\hat{w}_i}{\sigma_i}\right).$$

We therefore obtain for $i = 2, \ldots, N-2$, at iteration k

$$\begin{aligned}\hat{w}_i^{(k)}(s) &= \hat{w}_i^{(k-1)}(s) - \theta\left(\hat{\psi}_i^{(k)}(x_i, s) + \hat{\psi}_{i+1}^{(k)}(x_i, s)\right)\\ &= \hat{w}_i^{(k-1)}(s) - \theta\left(C_i\gamma_i + D_i + C_{i+1} + D_{i+1}\gamma_{i+1}\right).\end{aligned}$$

Using the identity $\gamma_i^2 - 1 = \sigma_i^2$ and simplifying, we get

(5.8)
$$\hat{w}_i^{(k)} = \hat{w}_i^{(k-1)} - \theta\left(\hat{w}_i^{(k-1)}\left(2 + \frac{2\gamma_i\gamma_{i+1}}{\sigma_i\sigma_{i+1}}\right) + \frac{\hat{w}_{i+1}^{(k-1)}}{\sigma_{i+1}}\left(\frac{\gamma_{i+2}}{\sigma_{i+2}} - \frac{\gamma_i}{\sigma_i}\right) + \frac{\hat{w}_{i-1}^{(k-1)}}{\sigma_i}\left(\frac{\gamma_{i-1}}{\sigma_{i-1}} - \frac{\gamma_{i+1}}{\sigma_{i+1}}\right) - \frac{\hat{w}_{i+2}^{(k-1)}}{\sigma_{i+1}\sigma_{i+2}} - \frac{\hat{w}_{i-2}^{(k-1)}}{\sigma_i\sigma_{i-1}}\right).$$

Substituting $\theta = 1/4$ yields the formula (5.6). For $i = 1$ and $i = N$, the Neumann conditions on the physical boundary are replaced by homogeneous Dirichlet conditions $\psi_1(0,t) = 0$ and $\psi_N(L,t) = 0$, $t > 0$. For these two subdomains, we obtain as solution after a Laplace transform

$$\hat{\psi}_1(x,s) = \frac{1}{\gamma_1}\left(\hat{w}_1\left(\frac{\gamma_1}{\sigma_1}+\frac{\gamma_2}{\sigma_2}\right) - \frac{\hat{w}_2}{\sigma_2}\right)\sinh\left((x-x_0)\sqrt{s}\right),$$
$$\hat{\psi}_N(x,s) = \frac{1}{\gamma_N}\left(\hat{w}_{N-1}\left(\frac{\gamma_{N-1}}{\sigma_{N-1}}+\frac{\gamma_N}{\sigma_N}\right) - \frac{\hat{w}_{N-2}}{\sigma_{N-1}}\right)\sinh\left((x_N-x)\sqrt{s}\right),$$

and thus the recurrence relation on the first interface is

(5.9)
$$\hat{w}_1^{(k)} = \hat{w}_1^{(k-1)} - \theta\left(\hat{w}_1^{(k-1)}\left(2+\frac{\gamma_1\gamma_2}{\sigma_1\sigma_2}+\frac{\sigma_1\gamma_2}{\gamma_1\sigma_2}\right) + \frac{\hat{w}_2^{(k-1)}}{\sigma_2}\left(\frac{\gamma_3}{\sigma_3}-\frac{\sigma_1}{\gamma_1}\right) - \frac{\hat{w}_3^{(k-1)}}{\sigma_2\sigma_3}\right).$$

On the last interface, we obtain

(5.10)
$$\begin{aligned}\hat{w}_{N-1}^{(k)} \quad = \quad & \hat{w}_{N-1}^{(k-1)} - \theta\left(\hat{w}_{N-1}^{(k-1)}\left(2+\frac{\gamma_{N-1}\gamma_N}{\sigma_{N-1}\sigma_N}+\frac{\sigma_N\gamma_{N-1}}{\gamma_N\sigma_{N-1}}\right)\right.\\ & \left. + \frac{\hat{w}_{N-2}^{(k-1)}}{\sigma_{N-1}}\left(\frac{\gamma_{N-2}}{\sigma_{N-2}}-\frac{\sigma_N}{\gamma_N}\right) - \frac{\hat{w}_{N-3}^{(k-1)}}{\sigma_{N-1}\sigma_{N-2}}\right).\end{aligned}$$

Substituting $\theta = 1/4$ into (5.9) and (5.10) leads to the required formulas (5.5) and (5.7), respectively. □

Lemma 5.1 shows that the interface traces at the kth iteration satisfy $\hat{\mathbf{w}}^{(k)} = A^k(s)\hat{\mathbf{w}}^{(0)}$. To prove superlinear convergence, we use a trick that is similar to the one in Theorem 2.3; that is, we write the recurrence as

(5.11)
$$\hat{\mathbf{w}}^{(k)} = M^k(s)\frac{\hat{\mathbf{w}}^{(0)}}{\sinh^{2k}(h_{\min}\sqrt{s})},$$

where $h_{\min} = \min_i h_i$ and $M(s) = \sinh^2(h_{\min}\sqrt{s})A(s)$. We see that the term $\frac{1}{\sinh^{2k}(h_{\min}\sqrt{s})}$ provides the necessary superlinear convergence, provided the L^1 norm of $\mathcal{L}^{-1}(M^k(s))$ grows only geometrically. This last fact is proved in the following lemma.

LEMMA 5.2. *Let $\{\nu_i^{(k)}(t)\}_{i=1}^{N-1}$ be functions whose Laplace transforms $\hat{\boldsymbol{\nu}}^{(k)}$ satisfy*

(5.12)
$$\hat{\boldsymbol{\nu}}^{(k)} = M(s)\hat{\boldsymbol{\nu}}^{(k-1)},$$

with $M(s) = \sinh^2(h_{\min}\sqrt{s})A(s)$ and $A(s)$ as defined in Lemma 5.1. Then there are constants $c_{ij} \geq 0$ such that $\sum_{j=1}^{N-1} c_{ij} \leq \frac{3}{2}$ and

(5.13)
$$\|\nu_i^{(k)}\|_{L^\infty(0,T)} \leq \sum_{j=1}^{N-1} c_{ij}\|\nu_j^{(k-1)}\|_{L^\infty(0,T)}, \quad k = 1,2,3,\ldots.$$

The analysis involves taking each entry of $M(s)$ and bounding the norm of its inverse Laplace transform. The calculations, which are somewhat cumbersome, are presented in Appendix A for completeness. Applying (5.13) k times shows that there are coefficients $c_{ij}^{(k)} \geq 0$ such that

(5.14)
$$\|\nu_i^{(k)}\|_{L^\infty(0,T)} \leq \sum_{j=1}^{N-1} c_{ij}^{(k)}\|\nu_j^{(0)}\|_{L^\infty(0,T)}$$

with $\sum_{j=1}^{N-1} c_{ij}^{(k)} \leq \left(\frac{3}{2}\right)^k$. This allows us to deduce the following theorem.

THEOREM 5.3. *Consider the NNWR algorithm (3.4)–(3.5) with initial guess $\{w_i^{(0)}(t)\}_{i=1}^{N-1}$ along the interfaces $\{x_i\}_{i=1}^{N-1}$. Then the interface values at the kth iteration $w_i^{(k)}(t)$ satisfy*

$$w_i^{(k)}(t) = \sum_{j=1}^{N-1} \int_0^t a_{ij}^{(k)}(\tau) w_j^{(0)}(t-\tau)\, d\tau, \tag{5.15}$$

where each $a_{ij}^{(k)}(t)$ is an infinitely differentiable function with $\frac{d^p}{dt^p} a_{ij}^{(k)}(0) = 0$ for all $p \geq 0$. Moreover, we have

$$\int_0^T |a_{ij}^{(k)}(\tau)|\, d\tau \leq c_{ij}^{(k)} B(h_{\min}, 2k, T), \tag{5.16}$$

where $B(\alpha, k, t)$ is defined in (4.5) and the constants $c_{ij}^{(k)} \geq 0$ satisfy $\sum_{j=1}^{N-1} c_{ij}^{(k)} \leq (3/2)^k$ for all i.

Proof. Equation (5.15) is just the inverse Laplace transform of the relation

$$\hat{\mathbf{w}}^{(k)} = A^k(s)\hat{\mathbf{w}}^{(0)},$$

where $a_{ij}^{(k)}(t)$ is the inverse Laplace transform of the (i,j)th entry of $A^k(s)$. Another way of writing $\hat{\mathbf{w}}^{(k)}$ is

$$\hat{\mathbf{w}}^{(k)} = \hat{\boldsymbol{\nu}}^{(k)} = M^k(s)\hat{\boldsymbol{\nu}}^{(0)}, \qquad \text{where} \quad \hat{\boldsymbol{\nu}}^{(0)} = \frac{\hat{\mathbf{w}}^{(0)}}{\sinh^{2k}(h_{\min}\sqrt{s})}.$$

The inequality (5.14) then implies

$$\|w_i^{(k)}\|_{L^\infty(0,T)} \leq \sum_{j=1}^{N-1} c_{ij}^{(k)} \|\phi_{2k} * w_j^{(0)}\|_{L^\infty(0,T)} \leq \sum_{j=1}^{N-1} c_{ij}^{(k)} \|w_j^{(0)}\|_{L^\infty(0,T)} \int_0^T |\phi_{2k}(\tau)|\, d\tau,$$

where $\phi_{2k}(t) = \mathcal{L}^{-1}\left(\frac{1}{\sinh^{2k}(h_{\min}\sqrt{s})}\right)$. Using the estimate in Lemma 4.5, we obtain the bound

$$\|w_i^{(k)}\|_{L^\infty(0,T)} \leq \sum_{j=1}^{N-1} c_{ij}^{(k)} B(h_{\min}, 2k, T) \|w_j^{(0)}\|_{L^\infty(0,T)},$$

which can only be true for all choices of w_j^0 if (5.16) holds. Finally, we note that $\hat{a}_{ij}^{(k)}$, the (i,j)th entry of $(A(s))^k$, is related to $\hat{\mu}_{ij}^{(k)}$, the (i,j)th entry of $(M(s))^k$, by

$$\hat{a}_{ij}^{(k)} = \frac{\hat{\mu}_{ij}^{(k)}}{\sinh^{2k}(h_{\min}\sqrt{s})}.$$

Since we have shown in Lemma 5.2 that the entries of $M(s)$ are either constants or Laplace transforms of regular functions, we deduce that

$$\lim_{s\to+\infty} s^p \hat{a}_{ij}^{(k)} = \lim_{s\to+\infty} \frac{s^p \hat{\mu}_{ij}^{(k)}}{\sinh^{2k}(h_{\min}\sqrt{s})} = 0.$$

Thus, derivatives of $a_{ij}^{(k)}(t)$ of all orders vanish at $t = 0$, as claimed. ∎

Proof of Theorem 3.1: From (5.15), we see that for all $1 \le i \le N-1$, we have

$$\begin{aligned}\|w_i^{(k)}\|_{L^\infty(0,T)} &\le \max_{1\le j\le N-1} \|w_j^{(0)}\|_{L^\infty(0,T)} \sum_{j=1}^{N-1} \int_0^T |a_{ij}^{(k)}(\tau)|\, d\tau \\ &\le \max_{1\le j\le N-1} \|w_j^{(0)}\|_{L^\infty(0,T)} \left(\frac{3}{2}\right)^k B(h_{\min}, 2k, T).\end{aligned}$$

By Lemma 4.5, we can estimate the right-hand side by

$$\|w_i^{(k)}\|_{L^\infty(0,T)} \le \max_{1\le j\le N-1} \|w_j^{(0)}\|_{L^\infty(0,T)} \left(\frac{3}{2}\right)^k \left(\frac{2}{1-e^{-(2k+1)h_{\min}^2/T}}\right)^{2k} e^{-k^2 h_{\min}^2/T}.$$

The estimate (3.6) in Theorem 3.1 now follows readily.

6. Analysis of the NNWR in 2D. In this section we formulate and analyze the NNWR algorithm, applied to the two-dimensional heat equation

$$\partial_t u - \Delta u = f(x,y,t), \quad (x,y) \in \Omega = (l,L)\times(0,\pi),\ t \in (0,T]$$

with initial condition $u(x,y,0) = u_0(x,y)$ and Dirichlet boundary conditions. To define the Neumann-Neumann algorithm, we decompose Ω into strips of the form $\Omega_i = (x_{i-1}, x_i) \times (0,\pi)$, $l = x_0 < x_1 < \cdots < x_N = L$. The Neumann-Neumann algorithm, considering directly the error equations with $f(x,y,t) = 0, u_0(x,y) = 0$ and homogeneous Dirichlet boundary conditions, is then given by performing iteratively for $k = 1, 2, \ldots$ and for $i = 1, \ldots, N$ the Dirichlet and Neumann steps

(6.1)

$$\begin{aligned}
\partial_t u_i^{(k)} - \Delta u_i^{(k)} &= 0, \quad \text{in } \Omega_i, & \partial_t \psi_i^{(k)} - \Delta \psi_i^{(k)} &= 0, \quad \text{in } \Omega_i,\\
u_i^{(k)}(x,y,0) &= 0, & \psi_i^{(k)}(x,y,0) &= 0,\\
u_i^{(k)}(x_{i-1},y,t) &= w_{i-1}^{(k-1)}(y,t), & \partial_{n_i}\psi_i^{(k)}(x_{i-1},y,t) &= (\partial_{n_{i-1}} u_{i-1}^{(k)} + \partial_{n_i} u_i^{(k)})(x_{i-1},y,t),\\
u_i^{(k)}(x_i,y,t) &= w_i^{(k-1)}(y,t), & \partial_{n_i}\psi_i^{(k)}(x_i,y,t) &= (\partial_{n_{i-1}} u_{i-1}^{(k)} + \partial_{n_i} u_i^{(k)})(x_i,y,t),\\
u_i^{(k)}(x,0,t) &= u_i^{(k)}(x,\pi,t) = 0, & \psi_i^{(k)}(x,0,t) &= \psi_i^{(k)}(x,\pi,t) = 0,
\end{aligned}$$

except for the first and last subdomain, where in the Neumann step the Neumann conditions are replaced by homogeneous Dirichlet conditions along the physical boundaries, as in the one dimensional case. The new interface values for the next step are then defined as

$$w_i^{(k)}(y,t) = w_i^{(k-1)}(y,t) - \theta\left(\psi_i^{(k)}(x_i,y,t) + \psi_{i+1}^{(k)}(x_i,y,t)\right),$$

with $\theta \in (0,1]$. Before stating the convergence result, we need two technical lemmas.

LEMMA 6.1. *Let $a_{ij}^{(k)} : [0,\infty) \to \mathbb{R}$ be defined as in Theorem 5.3. Then for each fixed $T > 0$, the functions $K_{ij}(t) := \sum_{n=1}^{\infty} a_{ij}^{(k)}(t) e^{-n^2 t}$ are uniformly bounded on the interval $[0,T]$.*

Proof. Since the $a_{ij}^{(k)}(t)$ are continuous and thus bounded on the interval $[0,T]$, the superlinearly decaying term $e^{-n^2 t}$ ensures convergence of the infinite sum for t bounded away from zero. Therefore, it suffices to show that $K_{ij}(t)$ remains bounded for t close to zero. Indeed, since the derivatives of all orders of $a_{ij}^{(k)}$ vanish at $t = 0$, Taylor's theorem

ensures the existence of a constant C such that $|a_{ij}^{(k)}(t)| \leq Ct^2$ for t small enough, so we have

$$\left| \sum_{n=1}^{M} a_{ij}^{(k)}(t) e^{-n^2 t} \right| \leq \frac{Ct^2}{1 - e^{-t}}. \tag{6.2}$$

In particular, for $0 < t < 1$, we have $1 - e^{-t} \geq t - \frac{t^2}{2} \geq \frac{t}{2}$, so the sum (6.2) is bounded above by $2Ct$, which is independent of M. Therefore, $K_{ij}(t)$ is bounded uniformly for all $t \in [0, T]$, as required. □

LEMMA 6.2. *Let $g \in L^\infty(0, \pi)$, and $\tau > 0$. Then for any $y \in \mathbb{R}$,*

$$\frac{2}{\pi} \left| \int_0^\pi g(\eta) \sum_{n \geq 1} e^{-n^2 \tau} \sin(n\eta) \sin(ny)\, d\eta \right| \leq \|g\|_{L^\infty(0,\pi)}.$$

Proof. Our first step is to rewrite the infinite sum using the Poisson summation formula, which we briefly recall here. Let $f \in L^2(\mathbb{R})$ and

$$\hat{f}(w) = \mathcal{F}(f(x))(w) := \int_{-\infty}^{\infty} f(t) e^{-iwt} dt$$

be its Fourier transform. Then

1. If both f and $\hat{f}$ are continuous and decay sufficiently rapidly, then
$$\textstyle\sum_{n \in \mathbb{Z}} f(n) = \sum_{k \in \mathbb{Z}} \hat{f}(2k\pi) \qquad \text{(Poisson summation formula)},$$
2. $\mathcal{F}\left(f(x) e^{iw_0 x}\right) = \hat{f}(w - w_0)$,
3. $\mathcal{F}\left(e^{-x^2 \tau}\right) = \sqrt{\frac{\pi}{\tau}} e^{-w^2/4\tau}$.

Since we can rewrite the infinite sum as

$$\begin{aligned} \frac{2}{\pi} \sum_{n \geq 1} e^{-n^2 \tau} \sin(n\eta) \sin(ny) &= \frac{1}{\pi} \sum_{n \geq 1} e^{-n^2 \tau} (\cos(n(\eta - y)) - \cos(n(\eta + y))) \\ &= \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{-n^2 \tau} (e^{in(\eta - y)} - e^{in(\eta + y)}), \end{aligned}$$

an application of the above properties and the Poisson summation formula gives

$$\frac{2}{\pi} \sum_{n \geq 1} e^{-n^2 \tau} \sin(n\eta) \sin(ny) = \frac{1}{\sqrt{4\pi\tau}} \sum_{k \in \mathbb{Z}} \left(e^{-(2k\pi - \eta + y)^2/4\tau} - e^{-(2k\pi - \eta - y)^2/4\tau} \right). \tag{6.3}$$

Next, let

$$I = \frac{2}{\pi} \int_0^\pi g(\eta) \sum_{n \geq 1} e^{-n^2 \tau} \sin(n\eta) \sin(ny)\, d\eta.$$

Noting that the sum on the right of (6.3) converges uniformly in η for $\eta \in [0, \pi]$, we can substitute it into the definition of I and exchange the integral and sum to get

$$\begin{aligned} I &= \frac{1}{\sqrt{4\pi\tau}} \left[\sum_{k \in \mathbb{Z}} \int_0^\pi g(\eta) e^{-\frac{(2k\pi - \eta + y)^2}{4\tau}} d\eta - \sum_{k \in \mathbb{Z}} \int_0^\pi g(\eta) e^{-\frac{(2k\pi - \eta - y)^2}{4\tau}} d\eta \right] \\ &= \frac{1}{\sqrt{4\pi\tau}} \left[\sum_{k \in \mathbb{Z}} \int_0^\pi g(\eta) e^{-\frac{(y - 2k\pi - \eta)^2}{4\tau}} d\eta - \sum_{k \in \mathbb{Z}} \int_0^\pi g(\eta) e^{-\frac{(y - 2k\pi + \eta)^2}{4\tau}} d\eta \right] \end{aligned}$$

We now perform the change of variables $\zeta = \eta + 2k\pi$ in the first integral and $\zeta = 2k\pi - \eta$ in the second integral to obtain

$$I = \frac{1}{\sqrt{4\pi\tau}} \left[\sum_{k\in\mathbb{Z}} \int_{2k\pi}^{(2k+1)\pi} g(\zeta - 2k\pi) e^{-\frac{(y-\zeta)^2}{4\tau}}\, d\zeta - \sum_{k\in\mathbb{Z}} \int_{(2k-1)\pi}^{2k\pi} g(2k\pi - \zeta) e^{-\frac{(y-\zeta)^2}{4\tau}}\, d\zeta \right].$$

Note that the integrals above are all taken over disjoint intervals. This motivates us to define the 2π-periodic odd extension of g in $\mathbb{R}$,

$$\bar{g}(y) = \begin{cases} g(y - 2k\pi), & 2k\pi \le y < (2k+1)\pi, \\ -g(2k\pi - y), & (2k-1)\pi \le y < 2k\pi, \quad (k \in \mathbb{Z}) \end{cases}$$

which allows us to write $I = \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} \bar{g}(\zeta) e^{-(y-\zeta)^2/4\tau}\, d\tau$. Finally, noting that $\bar{g}(y)$ has the same maxima and minima as $g(y)$, we deduce that

$$|I| \le \frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} |\bar{g}(\zeta)| e^{-\frac{(y-\zeta)^2}{4\tau}}\, d\zeta \le \|g\|_{L^\infty(0,\pi)} \underbrace{\frac{1}{\sqrt{4\pi\tau}} \int_{-\infty}^{\infty} e^{-\frac{(y-\zeta)^2}{4\tau}}\, d\zeta}_{=1} = \|g\|_{L^\infty(0,\pi)},$$

as required. ∎

We are now ready to prove the main result of this section, which states that we have the same convergence estimate as in the one-dimensional case.

THEOREM 6.3 (Convergence of NNWR in 2D). *Let $\theta = 1/4$. For $T > 0$ fixed, the NNWR algorithm (6.1) converges superlinearly with the estimate*

$$\max_{1\le i\le N-1} \|w_i^{(k)}\| \le \left(\frac{\sqrt{6}}{1 - e^{-\frac{(2k+1)h_{\min}^2}{T}}} \right)^{2k} e^{-k^2 h_{\min}^2/T} \max_{1\le i\le N-1} \|w_i^{(0)}\|, \tag{6.4}$$

where $h_{\min}$ is the minimum subdomain width and the norm $\|\cdot\|$ is $\|\cdot\|_{L^\infty(0,T;L^\infty(0,\pi))}$.

Proof. We reduce the 2D problem to a collection of one-dimensional problems by performing a Fourier sine transform along the y direction. Let $U_i^{(k)}(x,n,t)$ be the Fourier sine coefficients of $u_i^{(k)}(x,y,t)$, i.e.,

$$U_i^{(k)}(x,n,t) = \frac{2}{\pi} \int_0^\pi u_i^{(k)}(x,\eta,t) \sin(n\eta) d\eta.$$

Similarly, we define $W_i^{(k)}(n,t)$ to be the Fourier sine coefficients of the interface values $w_i^{(k)}(y,t)$. A direct calculation shows that the $U_i^{(k)}(x,n,t)$ satisfy

$$\frac{\partial U_i^{(k)}}{\partial t}(x,n,t) - \frac{\partial^2 U_i^{(k)}}{\partial x^2}(x,n,t) + n^2 U_i^{(k)}(x,n,t) = 0 \tag{6.5}$$

with the same boundary conditions as the one-dimensional case for each n. Taking Laplace transforms in t of (6.5) yields

$$(s+n^2)\hat{U}_i^{(k)} - \frac{d^2\hat{U}_i^{(k)}}{dx^2} = 0,$$

which is the same equation as the one analyzed in the proof of Theorem 3.1, except s is replaced by $s+n^2$. By mimicking the proof of Theorem 3.1, we see that the Laplace transform

$\hat{W}_i^{(k)}(n,s)$ of $W_i^{(k)}(n,t)$ satisfies

$$\hat{W}_i^{(k)}(n,s) = \sum_{j=1}^{N-1} \hat{a}_{ij}^{(k)}(s+n^2)\hat{W}_j^{(0)}(n,s), \tag{6.6}$$

where $\hat{a}_{ij}^{(k)}(s)$ is the Laplace transform of $a_{ij}^{(k)}(t)$ defined in Theorem 5.3. Taking the inverse Laplace transform of (6.6) and using Lemma 4.3, we obtain the time-domain equality

$$W_i^{(k)}(n,t) = \sum_{j=1}^{N-1} \int_0^t a_{ij}^{(k)}(\tau) e^{-n^2\tau} W_j^{(0)}(n,t-\tau)d\tau.$$

So the interface functions $w_i^{(k)}(y,t)$ can be written as

$$\begin{aligned} w_i^{(k)}(y,t) &= \sum_{n\geq 1} W_i^{(k)}(n,t)\sin(ny) \\ &= \sum_{n\geq 1}\sum_{j=1}^{N-1} \int_0^t a_{ij}^{(k)}(\tau)e^{-n^2\tau}\left(\frac{2}{\pi}\int_0^\pi w_j^{(0)}(\eta,t-\tau)\sin(n\eta)d\eta\right)\sin(ny)d\tau. \end{aligned}$$

Since $|\sum_{n\geq 1} a_{ij}^{(k)}(\tau)e^{-n^2\tau}|$ is uniformly bounded by Lemma 6.1, Fubini's theorem allows us to exchange the infinite sum and integral, so we get

$$w_i^{(k)}(y,t) = \sum_{j=1}^{N-1} \int_0^t a_{ij}^{(k)}(\tau)\left(\frac{2}{\pi}\int_0^\pi w_j^{(0)}(\eta,t-\tau)\sum_{n\geq 1} e^{-n^2\tau}\sin(n\eta)\sin(ny)d\eta\right)d\tau. \tag{6.7}$$

Lemma 6.2 now implies

$$\begin{aligned} |w_i^{(k)}(y,t)| &\leq \sum_{j=1}^{N-1}\int_0^t |a_{ij}^{(k)}(\tau)|\,\|w_j^{(0)}(\cdot,t-\tau)\|_{L^\infty(0,\pi)}\,d\tau \\ &\leq \max_{1\leq j\leq N-1}\|w_j^{(0)}\|_{L^\infty(0,t;L^\infty(0,\pi))}\sum_{j=1}^{N-1}\int_0^t |a_{ij}^{(k)}(\tau)|\,d\tau \\ &\leq \max_{1\leq j\leq N-1}\|w_j^{(0)}\|_{L^\infty(0,t;L^\infty(0,\pi))}\left(\frac{3}{2}\right)^k B(h_{\min},2k,t) \quad \text{(by Theorem 5.3).} \end{aligned}$$

Substituting the definition of $B(h_{\min},2k,t)$ and maximizing over $t\in[0,T]$ yields the bound (6.4), as required. □

7. Numerical Experiments. We perform experiments to measure the actual convergence rate of the discretized DNWR and NNWR algorithms for the problem

$$\begin{array}{rcll} \partial_t u - \frac{\partial}{\partial x}(\kappa(x)\partial_x u) &=& 0, & x\in\Omega, \\ u(x,0) &=& x(x+1)(x+3)(x-2)e^{-x}, & x\in\Omega, \\ u(-3,t)=t,\ u(2,t) &=& te^{-t}, & t>0. \end{array} \tag{7.1}$$

We discretize (7.1) using standard centered finite differences in space and backward Euler in time on a grid with $\Delta x = 2\times 10^{-2}$ and $\Delta t = 4\times 10^{-3}$. Note that in some of the experiments below, the diffusion coefficient $\kappa(x)$ will be spatially varying. This will allow us to study how spatially varying coefficients affect the convergence of our algorithms, which have only been analyzed in the constant coefficient case.

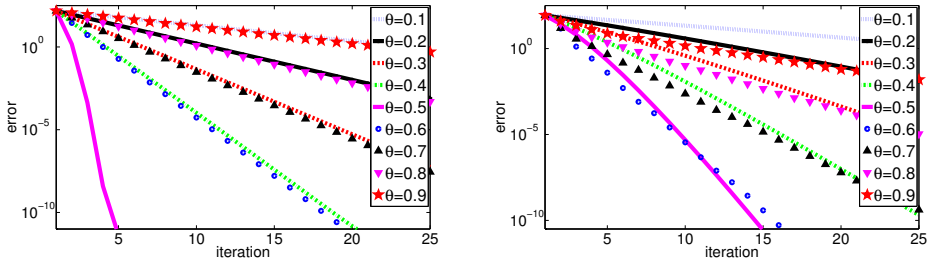

FIG. 7.1. *Convergence of DNWR for $a > b$ using various relaxation parameters θ for $T = 2$, on the left for $\kappa(x) = 1$ and on the right for $\kappa(x) = 1 + e^x$*

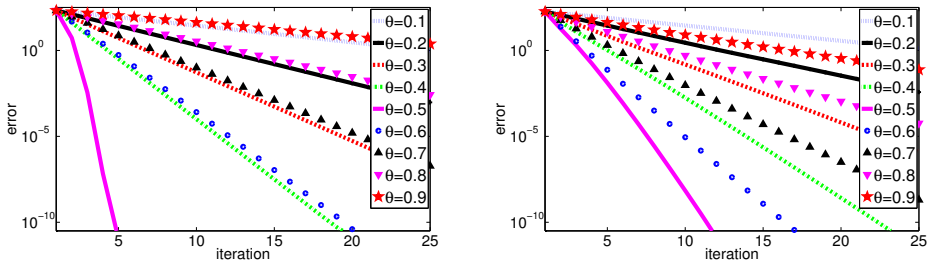

FIG. 7.2. *Convergence of DNWR for $a < b$ using various relaxation parameters θ for $T = 2$, on the left for $\kappa(x) = 1$ and on the right for $\kappa(x) = 1 + e^x$*

7.1. The DNWR method. Our first set of experiments will illustrate the DNWR method. We consider the spatial domain $\Omega := (-3, 2)$ and split it into two non-overlapping subdomains $\Omega_1 = (-3, 0)$ and $\Omega_2 = (0, 2)$. In the first case, we choose the larger domain to be the Dirichlet subdomain, and the smaller one to be the Neumann subdomain; this corresponds to the case of $a = 3$, $b = 2$, $a > b$ in Theorem 2.2. We then run the DNWR algorithm for the time window $T = 2$ with initial guess $w^{(0)}(t) = t^2, t \in [0, T]$: Figure 7.1 shows the convergence behavior for various values of θ and for two choices of diffusion coefficients, $\kappa(x) = 1$ on the left and $\kappa(x) = 1 + e^x$ on the right. We see that for this reasonably small time window, we get linear convergence for all relaxation parameters θ, except for $\theta = 1/2$, when we observe superlinear convergence. This is the case regardless of whether the diffusion coefficient varies spatially or not.

In the second case, we repeat the experiment, except we swap the roles of the two subdomains, so that the Dirichlet domain is now smaller than the Neumann one. This corresponds to $a = 2$ and $b = 3$, as in Theorem 2.3. The diffusion coefficients and initial guess are the same as before, and the results are shown in Figure 7.2. We see once again that $\theta = 1/2$ is the choice that gives superlinear (and often the fastest) convergence, for both the constant and spatially-varying diffusion coefficient.

7.2. The NNWR method. Next, we show an experiment for the NNWR algorithm in the spatial domain $\Omega = (0, 6)$, with the same discretization parameters Δx and Δt as above, and for the time window $T = 2$. From now on, we always consider $\kappa(x) = 1$, unless otherwise specified. In Figure 7.3, we consider a decomposition into two to six unequal subdomains, whose widths are shown in Table 7.1. On the left panel, we show the convergence in the four-subdomain case as a function of the relaxation parameter θ, whereas on the right panel, we show the convergence for $\theta = 1/4$ as we vary the number of subdomains.

TABLE 7.1
Subdomain lengths used for the NNWR experiments in Fig. 7.3.

No. of subdomains	h_1	h_2	h_3	h_4	h_5	h_6
2	3.50	2.50				
3	2.30	2.30	1.40			
4	1.20	2.40	1.80	0.60		
5	1.80	1.40	1.08	1.00	0.72	
6	1.20	0.80	1.00	1.20	1.00	0.80

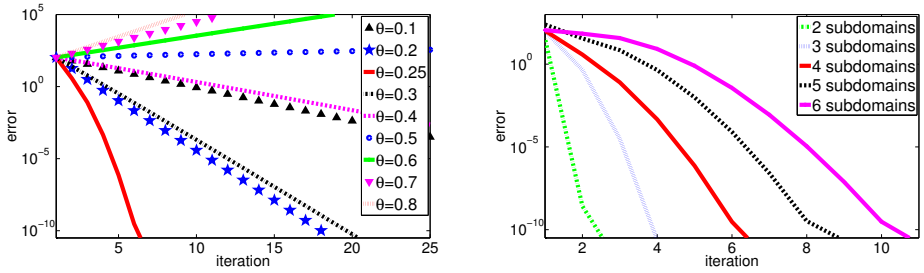

FIG. 7.3. *Convergence of NNWR with four subdomains and various relaxation parameters on the left, and dependence of NNWR on the number of subdomains for $\theta = 1/4$ on the right*

We observe superlinear convergence for $\theta = 1/4$, and only linear convergence for the other choices. We also see that convergence slows down as the number of subdomains is increased, as expected.

We now compare the numerical behavior of DNWR and NNWR with our theoretical estimates in Sections 2 and 3. In Figure 7.4, we show for the DNWR algorithm a comparison between the numerically measured convergence for the discretized problem, the theoretical convergence for the continuous model problem (calculated using inverse Laplace transforms), and the linear and superlinear convergence estimates shown in Theorem 2.2, for $a = 3$, $b = 2$, $\kappa(x) = 1$. We see that for a short time interval, $T = 2$, the algorithm converges superlinearly, and the superlinear bound is quite accurate. For the long time interval $T = 50$, the algorithm converges linearly, and the linear convergence estimate is now more accurate. Similarly, we show in Figure 7.5 a comparison of the numerically measured convergence for the NNWR algorithm for $\theta = 1/4$ and $\kappa(x) = 1$, and the theoretical estimates from Theorem 3.1. On the left, we show the results for the two subdomain case (subdomain lengths are as in the first line of Table 7.1), where we also plotted the linear estimate from [28], and on the right, we show the results for the case of many subdomains of equal length for $\Omega = (0, 6)$. We see that although the superlinear bounds are sometimes not sharp, we do manage to capture the superlinear nature of the convergence, which would be difficult with abstract Schwarz type techniques. We now compare in Figure 7.6 the performance of the DNWR and NNWR algorithms for two subdomains with the Schwarz Waveform Relaxation algorithms from [16, 2] with overlap. We use an overlap of length $2\Delta x$, where $\Delta x = 1/50$. We observe that the DNWR and NNWR algorithms converge faster than the overlapping Schwarz WR iteration. Only a higher order optimized Schwarz waveform relaxation algorithm comes close to the performance of the DNWR algorithm in this experiment.

7.3. An example with non-matching time steps. One of the advantages of WR is the ability to use different time step sizes for different subdomains. Here, we show results for the

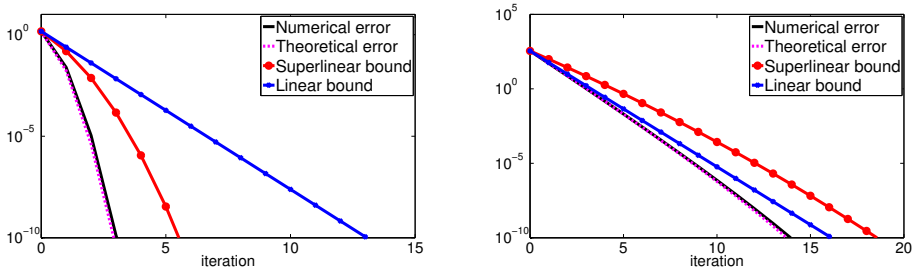

FIG. 7.4. *Comparison of the numerically measured convergence rates and the theoretical error estimates for DNWR for* $\kappa(x) = 1$ *with* $T = 2$ *on the left, and* $T = 50$ *on the right*

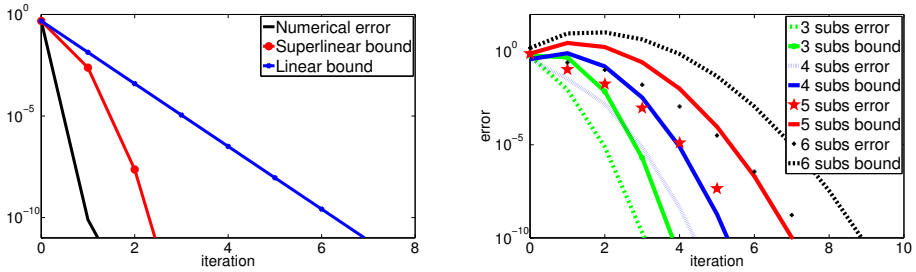

FIG. 7.5. *Comparison of the numerically measured convergence rates and the theoretical error estimates for NNWR for* $\kappa(x) = 1$ *with* $\theta = 1/4$ *and* $T = 2$*, on the left for two subdomains, and on the right for many subdomains*

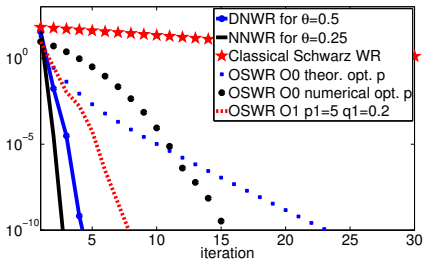

FIG. 7.6. *Comparison of DNWR and NNWR with Schwarz waveform relaxation*

TABLE 7.2
Time steps for different subdomains

	Ω_1	Ω_2	Ω_3
time grids Δt_i	3.8×10^{-3}	2×10^{-3}	7.4×10^{-3}

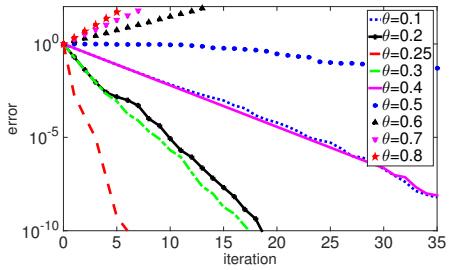

FIG. 7.7. *Convergence plot of the NNWR with three heterogeneous subdomains for various values of the relaxation parameter θ*

NNWR method with three equal subdomains applied to the model problem

$$u_t - u_{xx} = 0 \quad \text{on } \Omega \times (0,T) = (0,6) \times (0,1),$$
$$u(x,0) = 0, \quad u(0,t) = t^2, \quad u(6,t) = t^3.$$

For the spatial discretization, we use finite differences on a uniform mesh of size $\Delta x = 0.01$. For the time discretization, we use a uniform time grid Δt_i on Ω_i for $i = 1, 2, 3$, where the values of Δt_i are given in Table 7.2. We plot the error of the method for different values of θ as a function of the number of iterations. We see that $\theta = 0.25$ is the parameter that leads to the fastest convergence, just as predicted by the analysis at the continuous level.

7.4. Rothe's method versus NNWR. To illustrate the advantages of waveform relaxation over the traditional Rothe's method under the right circumstances, we consider solving the problem (7.1) for $T = 0.2$ and 4000 time steps ($\Delta t = 0.2/4000$) using a fixed number of processors $N = 24$. We compare three approaches:

A. NNWR, using N processors to compute in N subdomains,
B. Pipeline NNWR, using J processors *per subdomain* to compute in N/J subdomains,
C. Rothe's method with classical NN for the elliptic subproblem at each time step; we use N processors to compute in N subdomains.

The pipeline approach was proposed for Schwarz WR in [41] and can be extended to NNWR. In essence, each subdomain is assigned multiple processors, and the first processor for each subdomain starts integrating forward in time. After a fixed number of time steps, *but before the end of the time window*, interface data is already transmitted to the neighbours, so that the second iteration can already begin *while the first iteration is still on-going*. This can be repeated as long as there are additional processors available, so that several iterations can operate simultaneously on different parts of the time window in a pipelined fashion. An illustration of how the method works for two subdomains is shown in Figure 7.8. Note that the pipelined version is mathematically identical to the original NNWR; the only difference lies in the organization of the computation and communication. Tables 7.3 and 7.4 compare the three approaches mentioned above for two different spatial grid sizes ($\Delta x = 1/8000$

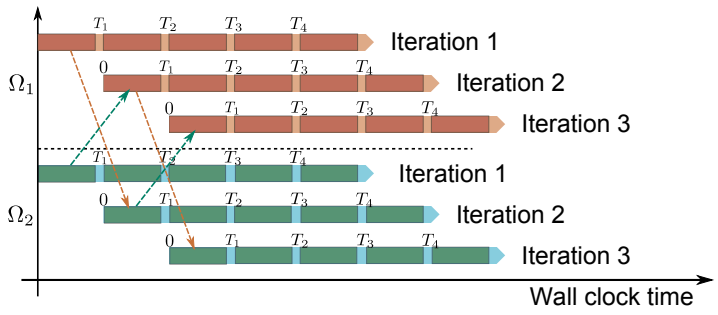

FIG. 7.8. *Pipeline waveform relaxation method using six processors for two subdomains. Each bar represents a processor, with the darker color indicating time spent on numerical integration, and the lighter color indicating time spent on communication with neighboring subdomains. The T_i above the bars show which portion of the time window is being worked on at a particular instant in wall-clock time. The arrows indicate the transfer of interface data between subdomains.*

TABLE 7.3
Comparison for $\Delta x = 1/8000$, $\Delta t = 0.2/4000$

Method	No. subdomains	Iterations	Total wall time (sec.)
A	24	10	7.59
B	12	4	3.82
C	24	2 per time step	5.76

TABLE 7.4
Comparison for $\Delta x = 1/16000$, $\Delta t = 0.2/4000$

Method	No. subdomains	Iterations	Total wall time (sec.)
A	24	10	11.39
B	12	4	5.05
B	6	3	3.43
C	24	2 per time step	7.11

and $\Delta x = 1/16000$). The total wall-clock time was obtained on the *Superior*, the high-performance computing cluster at Michigan Technological University. We see that without pipelining, the original NNWR method requires too many iterations to be competitive with Rothe's method. However, when pipelining is used, the iteration count decreases so much that communication becomes the dominant cost of the computation, and the total wall-clock time is now lower for pipeline NNWR than for Rothe's method. This is why it is important to understand the behavior of the NNWR method in detail.

7.5. A two-dimensional case. We show an experiment for the NNWR algorithm in two dimensions for the following model problem

$$\partial_t u - (\partial_{xx} u + \partial_{yy} u) = 0, \quad u(x, y, 0) = \sin(2\pi x) \sin(3y).$$

We decompose our domain $\Omega := (0, 1) \times (0, \pi)$ into three non-overlapping subdomains $\Omega_1 = (0, 2/5) \times (0, \pi)$, $\Omega_2 = (2/5, 3/4) \times (0, \pi)$, $\Omega_3 = (3/4, 1) \times (0, \pi)$, see Figure 7.9 on the left. On the right, we plot the numerical errors of the NNWR algorithm for various θ and the theoretical estimates from Theorem 6.3 for $\theta = 1/4$ and again observe superlinear convergence.

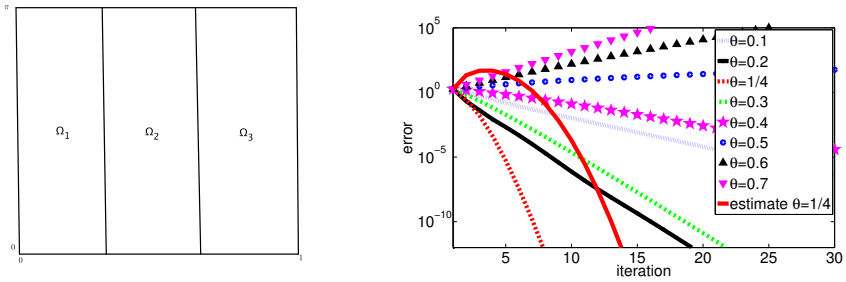

FIG. 7.9. *Decomposition of 2D domain into strips on the left, and convergence of NNWR using various relaxation parameters θ for $T = 0.2$ on the right*

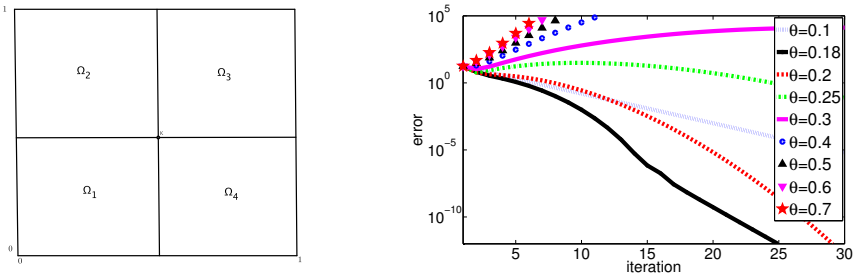

FIG. 7.10. *Decomposition of a 2D domain with a cross point on the left, and convergence of NNWR using various relaxation parameters θ for $T = 0.2$ on the right*

7.6. An example with a cross point. We conclude this subsection with a further numerical experiment not covered by our analysis: we decompose the two dimensional domain $\Omega := (0,1) \times (0,1)$ into four non-overlapping subdomains $\Omega_1 = (0,1/2) \times (0,1/2)$, $\Omega_2 = (0,1/2) \times (1/2,1)$, $\Omega_3 = (1/2,1) \times (1/2,1)$, $\Omega_4 = (1/2,1) \times (0,1/2)$, such that a cross point is present, see Figure 7.10 on the left. We take the initial condition as: $u(x,y,0) = \sin(2\pi x)\sin(3\pi y)$. The right panel of that figure shows that the convergence of the NNWR algorithm remains superlinear, despite the presence of the cross point.

8. Conclusions. We introduced two new classes of space-time parallel algorithms, the Dirichlet-Neumann waveform relaxation (DNWR) and the Neumann-Neumann waveform relaxation (NNWR) algorithms. For the one-dimensional heat equation, we proved superlinear convergence for both algorithms for a particular choice of the relaxation parameter. For the NNWR, our convergence estimate holds for a decomposition into many subdomains, and we also gave an extension to two spatial dimensions for a specific decomposition into strips. We are currently working on the generalization of our analysis for the DNWR algorithm to the case of many subdomains and higher spatial dimensions, and we are also studying the convergence for $\theta \neq 1/2$ in more detail.

Appendix A. Proof of Lemma 5.2. Define $\sigma := \sinh(h_{\min}\sqrt{s})$. By Lemma 5.1, the matrix $M(s) = (\hat{\mu}_{ij}(s))_{i,j=1}^{N-1}$ is pentadiagonal with entries

$$\hat{\mu}_{i,i} = -\frac{\sigma^2(\gamma_i\gamma_{i+1} - \sigma_i\sigma_{i+1})}{2\sigma_i\sigma_{i+1}}, \qquad \hat{\mu}_{i,i+1} = -\frac{\sigma^2(\sigma_i\gamma_{i+2} - \gamma_i\sigma_{i+2})}{4\sigma_i\sigma_{i+1}\sigma_{i+2}},$$

$$\hat{\mu}_{i,i-1} = -\frac{\sigma^2(\sigma_{i+1}\gamma_{i-1} - \gamma_{i+1}\sigma_{i-1})}{4\sigma_i\sigma_{i-1}\sigma_{i+1}}, \qquad \hat{\mu}_{i,i+2} = -\frac{\sigma^2}{4\sigma_{i+1}\sigma_{i+2}}, \qquad \hat{\mu}_{i,i-2} = -\frac{\sigma^2}{4\sigma_i\sigma_{i-1}}$$

for $i = 2, \dots, N-2$, and similar expressions for $i = 1, N-1$ (to be shown below). To deduce the inequality in (5.13), we show that each $\hat{\mu}_{ij}$ is either a constant or the Laplace transform of a regular function. In the latter case, we denote the inverse Laplace transform of $\hat{\mu}_{ij}(s)$ by $\mu_{ij}(t)$ (without the hat).

We first consider $\hat{\mu}_{i,i+2}$. If $h_{i+1} = h_{i+2} = h_{\min}$, then the terms in $\hat{\mu}_{i,i+2}$ cancel and we simply get $\hat{\mu}_{i,i+2} = -1/4$. If $h_{\min} \le h_{i+1}$ and $h_{\min} \le h_{i+2}$, then the kernel $\mu_{i,i+2}$, being a convolution of two non-negative functions, is non-negative by part 1 of Lemma 4.1, and using Lemma 4.2 its integral is bounded by

$$\int_0^\infty |\mu_{i,i+2}(t)|dt \le \lim_{s\to 0+} \frac{\sinh^2(h_{\min}\sqrt{s})}{4\sinh(h_{i+1}\sqrt{s})\sinh(h_{i+2}\sqrt{s})} \le \frac{h_{\min}^2}{4h_{i+1}h_{i+2}} \le \frac{1}{4}.$$

So in either case, we have

$$\left\| \mathcal{L}^{-1}\left(\hat{\mu}_{i,i+2}\,\hat{\nu}_{i+2}^{(k-1)}(s)\right)\right\|_{L^\infty(0,T)} \le \frac{1}{4}\|\nu_{i+2}^{(k-1)}\|_{L^\infty(0,T)}.$$

The same argument also holds for the term involving $\hat{\nu}_{i-2}^{(k-1)}$. Now for $\hat{\nu}_{i+1}^{(k-1)}$, we rewrite $\hat{\mu}_{i,i+1}$ as

$$\hat{\mu}_{i,i+1} = -\frac{\sinh\left((h_i - h_{i+2})\sqrt{s}\right)\sinh^2(h_{\min}\sqrt{s})}{4\sinh(h_i\sqrt{s})\sinh(h_{i+1}\sqrt{s})\sinh(h_{i+2}\sqrt{s})}.$$

We will again "pair" the factors in the numerator and denominator to prove that $\mu_{i,i+1}$ is a convolution of functions that do not change signs. If $h_i \ge h_{i+2}$, then $|h_i - h_{i+2}| < h_i$, so we use the pairing

$$\hat{\mu}_{i,i+1} = -\frac{\sinh((h_i - h_{i+2})\sqrt{s})}{4\sinh(h_i\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_{i+1}\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_{i+2}\sqrt{s})}.$$

On the other hand, if $h_i < h_{i+2}$, then $|h_i - h_{i+2}| < h_{i+2}$, so we use the pairing

$$\hat{\mu}_{i,i+1} = \frac{\sinh((h_{i+2} - h_i)\sqrt{s})}{4\sinh(h_{i+2}\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_{i+1}\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_i\sqrt{s})}.$$

In both cases, each factor is either a constant or the inverse Laplace transform of a non-negative function by Lemma 4.4, so the L^1 norm of $\mu_{i,i+1}$ can be estimated by taking the limit as $s \to 0+$:

$$\int_0^\infty |\mu_{i,i+1}(t)|dt \le \frac{|h_i - h_{i+2}|h_{\min}^2}{4h_ih_{i+1}h_{i+2}} < \frac{1}{4}.$$

This, in turn, implies $\|\mathcal{L}^{-1}(\hat{\mu}_{i,i+1}\hat{\nu}_{i+1}^{(k-1)}(s))\|_{L^\infty(0,T)} \le \frac{1}{4}\|\nu_{i+1}^{(k-1)}\|_{L^\infty(0,T)}$. A similar bound holds for the term involving $\hat{\nu}_{i-1}^{(k-1)}$.

Finally, for $\hat{\mu}_{i,i}$, we first note that

$$\hat{\mu}_{i,i} = -\frac{\sinh^2(h_{\min}\sqrt{s})\cosh((h_{i+1} - h_i)\sqrt{s})}{2\sinh(h_i\sqrt{s})\sinh(h_{i+1}\sqrt{s})}.$$

Using the identity $\sinh(u)\cosh(v) = \frac{1}{2}(\sinh(u-v) + \sinh(u+v))$, we obtain

$$\begin{aligned}\hat{\mu}_{i,i} = &-\frac{\sinh(h_{\min}\sqrt{s})\sinh\left((h_{\min} + h_i - h_{i+1})\sqrt{s}\right)}{4\sinh(h_i\sqrt{s})\sinh(h_{i+1}\sqrt{s})} \\ &-\frac{\sinh(h_{\min}\sqrt{s})\sinh\left((h_{\min} + h_{i+1} - h_i)\sqrt{s}\right)}{4\sinh(h_i\sqrt{s})\sinh(h_{i+1}\sqrt{s})}.\end{aligned}$$

Each term is again a ratio of hyperbolic sines, so we only need to pair the factors so that the coefficient in the numerator is always smaller than the one in the denominator. Now $-h_{i+1} \leq h_{\min} + h_i - h_{i+1} = h_i + h_{\min} - h_{i+1} \leq h_i$, so for the first term, we choose the pairing

$$-\frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_{i+1}\sqrt{s})} \cdot \frac{\sinh\left((h_{\min}+h_i-h_{i+1})\sqrt{s}\right)}{4\sinh(h_i\sqrt{s})} \quad \text{if } h_{i+1} \leq h_i,$$

and

$$-\frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_i\sqrt{s})} \cdot \frac{\sinh\left((h_{\min}+h_i-h_{i+1})\sqrt{s}\right)}{4\sinh(h_{i+1}\sqrt{s})} \quad \text{if } h_{i+1} \geq h_i.$$

A similar argument holds also for the second term. Now using Lemma 4.2, we again conclude that the integral of each kernel is bounded by $1/4$. In summary, we get for $2 \leq i \leq N-2$

$$\begin{aligned}\|\nu_i^{(k)}\|_{L^\infty(0,T)} \leq \frac{1}{2}\|\nu_i^{(k-1)}\|_{L^\infty(0,T)} + \frac{1}{4}\Big(&\|\nu_{i-2}^{(k-1)}\|_{L^\infty(0,T)} + \|\nu_{i-1}^{(k-1)}\|_{L^\infty(0,T)} \\ &+ \|\nu_{i+1}^{(k-1)}\|_{L^\infty(0,T)} + \|\nu_{i+2}^{(k-1)}\|_{L^\infty(0,T)}\Big),\end{aligned}$$

and the estimate (5.13) is established for interior subdomains. To complete the proof, we need to perform a similar analysis on

$$\hat{\mu}_{1,1} = -\frac{\sigma^2}{4}\left(\frac{\gamma_1\gamma_2-\sigma_1\sigma_2}{\sigma_1\sigma_2} + \frac{\sigma_1\gamma_2-\gamma_1\sigma_2}{\gamma_1\sigma_2}\right),\ \hat{\mu}_{1,2} = -\frac{\sigma^2}{4\sigma_2}\left(\frac{\gamma_3}{\sigma_3} - \frac{\sigma_1}{\gamma_1}\right),\ \hat{\mu}_{1,3} = \frac{1}{4\sigma_2\sigma_3}$$

and on the corresponding kernels for $i = N-1$, see (5.7). We only present the calculations for the first interface, the last one being similar. First, the kernel $\mu_{1,3}$ can be estimated like the interior kernels $\mu_{i,i+2}$. For $\hat{\mu}_{1,2}$, we have

$$\hat{\mu}_{1,2} = -\frac{\cosh\left((h_1-h_3)\sqrt{s}\right)\sinh^2(h_{\min}\sqrt{s})}{4\cosh(h_1\sqrt{s})\sinh(h_2\sqrt{s})\sinh(h_3\sqrt{s})}.$$

If $h_1 \geq h_3$, the decomposition

$$\hat{\mu}_{1,2} = -\frac{\cosh\left((h_1-h_3)\sqrt{s}\right)}{4\cosh(h_1\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_2\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_3\sqrt{s})}$$

shows that one can bound the L^1 norm of $\mu_{1,2}$ by

$$\int_0^\infty |\mu_{1,2}(t)|dt \leq \frac{h_{\min}^2}{4h_2h_3} < \frac{1}{4}.$$

If $h_3 > h_1$, then we rewrite

$$\begin{aligned}\hat{\mu}_{1,2} = -\frac{1}{4\cosh(h_1\sqrt{s})} \cdot \frac{\sinh(h_{\min}\sqrt{s})}{\sinh(h_2\sqrt{s})}\Bigg(&\frac{\sinh\left((h_{\min}+h_1-h_3)\sqrt{s}\right)}{2\sinh(h_3\sqrt{s})} \\ &+\frac{\sinh\left((h_{\min}+h_3-h_1)\sqrt{s}\right)}{2\sinh(h_3\sqrt{s})}\Bigg),\end{aligned}$$

which again shows, thanks to Lemma 4.2, that the integral is bounded by $1/4$. Finally we consider $\hat{\mu}_{1,1}$, whose expression can be manipulated to give

$$\begin{aligned}
\hat{\mu}_{1,1} &= -\frac{\sigma^2((\gamma_1^2+\sigma_1^2)\gamma_2 - 2\sigma_1\gamma_1\sigma_2)}{4\sigma_1\gamma_1\sigma_2} \\
&= -\frac{\sinh^2(h_{\min}\sqrt{s})(\cosh(2h_1\sqrt{s})\cosh(h_2\sqrt{s}) - \sinh(2h_1\sqrt{s})\sinh(h_2\sqrt{s}))}{2\sinh(2h_1\sqrt{s})\sinh(h_2\sqrt{s})} \\
&= -\frac{\cosh\left((2h_1-h_2)\sqrt{s}\right)\sinh^2(h_{\min}\sqrt{s})}{2\sinh(2h_1\sqrt{s})\sinh(h_2\sqrt{s})} \\
&= -\frac{\sinh(h_{\min}\sqrt{s})\sinh\left((h_{\min}+2h_1-h_2)\sqrt{s}\right)}{4\sinh(2h_1\sqrt{s})\sinh(h_2\sqrt{s})} \\
&\quad -\frac{\sinh(h_{\min}\sqrt{s})\sinh\left((h_{\min}-2h_1+h_2)\sqrt{s}\right)}{4\sinh(2h_1\sqrt{s})\sinh(h_2\sqrt{s})}.
\end{aligned}$$

Using the inequalities $-h_2 \le h_{\min} + 2h_1 - h_2 \le 2h_1$ and $-2h_1 \le h_{\min} - 2h_1 + h_2 \le h_2$ we can again, with an appropriate pairing of factors, bound the integral of each term by $1/4$. Thus, we have shown that for the subdomain touching the left physical boundary, we have the inequality

$$\|\nu_1^{(k)}\|_{L^\infty(0,T)} \le \frac{1}{2}\|\nu_1^{(k-1)}\|_{L^\infty(0,T)} + \frac{1}{4}\left(\|\nu_2^{(k-1)}\|_{L^\infty(0,T)} + \|\nu_3^{(k-1)}\|_{L^\infty(0,T)}\right).$$

A similar result holds for $\hat{\nu}_{N-1}^{(k)}(s)$, and hence the inequality (5.13) holds for all $1 \le i \le N-1$.

REFERENCES

[1] A. BELLEN AND M. ZENNARO, *The use of Runge-Kutta formulae in waveform relaxation methods*, Appl. Numer. Math., 11 (1993), pp. 95–114.

[2] D. BENNEQUIN, M. J. GANDER, AND L. HALPERN, *A Homographic Best Approximation Problem with Application to Optimized Schwarz Waveform Relaxation*, Math. of Comp., (2009), pp. 185–223.

[3] M. BJØRHUS, *A note on the convergence of discretized dynamic iteration*, BIT, (1995), pp. 291–296.

[4] P. E. BJØRSTAD AND O. B. WIDLUND, *Iterative Methods for the Solution of Elliptic Problems on Regions Partitioned into Substructures*, SIAM J. Numer. Anal., 23 (1986), pp. 1097–1120.

[5] J. F. BOURGAT, R. GLOWINSKI, P. LE TALLEC, AND M. VIDRASCU, *Variational Formulation and Algorithm for Trace Operator in Domain Decomposition Calculations*, in Domain Decomposition Methods, T. F. Chan, R. Glowinski, J. Périaux, and O. B. Widlund, eds., SIAM, 1989, pp. 3–16.

[6] J. H. BRAMBLE, J. E. PASCIAK, AND A. H. SCHATZ, *An Iterative Method for Elliptic Problems on Regions Partitioned into Substructures*, Math. Comp., 46 (1986), pp. 361–369.

[7] X.-C. CAI, *Additive Schwarz algorithms for parabolic convection-diffusion equations*, Numer. Math., 60 (1991), pp. 41–61.

[8] ———, *Multiplicative Schwarz Methods for Parabolic Problems*, SIAM J. Sci. Comput., 15 (1994), pp. 587–603.

[9] T. F. CHAN AND T. P. MATHEW, *Domain decomposition algorithms*, Acta numerica, 3 (1994), pp. 61–143.

[10] R. V. CHURCHILL, *Operational Mathematics*, McGraw-Hill, 2nd ed., 1958.

[11] Y. DE ROECK AND P. LE TALLEC, *Analysis and Test of a local domain decomposition preconditioner*, in Domain Decomposition Methods for PDEs, I, R. Glowinski et al., ed., Philadelphia, 1991, SIAM, pp. 112–128.

[12] C. FARHAT AND F.-X. ROUX, *A method of finite element tearing and interconnecting and its parallel solution algorithm*, Internat. J. Numer. Methods Engrg., 32 (1991), pp. 1205–1227.

[13] M. J. GANDER, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–732.

[14] M. J. GANDER, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition, T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, eds., Springer, 2015, pp. 69–114.

[15] M. J. GANDER AND L. HALPERN, *Absorbing Boundary Conditions for the Wave Equation and Parallel Computing*, Math. of Comput., 74 (2004), pp. 153–176.

[16] ———, *Optimized Schwarz Waveform Relaxation for Advection Reaction Diffusion Problems*, SIAM J. Numer. Anal., 45 (2007), pp. 666–697.
[17] M. J. GANDER, L. HALPERN, AND F. NATAF, *Optimal Schwarz Waveform Relaxation for the One Dimensional Wave Equation*, SIAM J. Num. Anal., 41 (2003), pp. 1643–1681.
[18] M. J. GANDER, F. KWOK, AND B. C. MANDAL, *Dirichlet-neumann waveform relaxation method for the 1d and 2d heat and wave equations in multiple subdomains*, Submitted, (2015). available at arXiv:1507.04011 [math.AP].
[19] M. J. GANDER AND A. M. STUART, *Space-time continuous analysis of waveform relaxation for the heat equation*, SIAM J. Sci. Comput., 19 (1998), pp. 2014–2031.
[20] M. J. GANDER AND H. ZHAO, *Overlapping Schwarz Waveform Relaxation for the Heat Equation in n-Dimensions*, BIT, 42 (2002), pp. 779–795.
[21] E. GILADI AND H. KELLER, *Space time domain decomposition for parabolic problems*, Tech. Report 97-4, Center for research on parallel computation CRPC, Caltech, 1997.
[22] L. HALPERN, C. JAPHET, AND J. SZEFTEL, *Optimized Schwarz waveform relaxation and discontinuous galerkin time stepping for heterogeneous problems*, SIAM J. Numer. Anal., 50 (2012), pp. 2588–2611.
[23] M. HEINKENSCHLOSS AND M. HERTY, *A spatial domain decomposition method for parabolic optimal control problems*, J. Comput. Appl. Math., 201 (2007), pp. 88–111.
[24] T. T. P. HOANG, *Space-time domain decomposition methods for mixed formulations of flow and transport problems in porous media*, PhD thesis, University Paris 6, France, 2013.
[25] T.-T. P. HOANG, J. JAFFRÉ, C. JAPHET, M. KERN, AND J. E. ROBERTS, *Space-time domain decomposition methods for diffusion problems in mixed formulations*, SIAM J. Numer. Anal., 51 (2013), pp. 3532–3559.
[26] G. HORTON AND S. VANDEWALLE, *A space-time multigrid method for parabolic partial differential equations*, SIAM Journal on Scientific Computing, 16 (1995), pp. 848–864.
[27] R. JELTSCH AND B. POHL, *Waveform relaxation with overlapping splittings*, SIAM J. Sci. Comput., 16 (1995), pp. 40–49.
[28] F. KWOK, *Neumann-Neumann Waveform Relaxation for the Time-Dependent Heat Equation*, in Domain Decomposition in Science and Engineering XXI, J. Erhel, M. J. Gander, L. Halpern, G. Pichot, T. Sassi, and O. B. Widlund, eds., vol. 98, Springer-Verlag, 2014, pp. 189–198.
[29] P. LE TALLEC, Y. DE ROECK, AND M. VIDRASCU, *Domain decomposition methods for large linearly elliptic three-dimensional problems*, J. of Comput. and App. Math., 34 (1991), pp. 93–117.
[30] E. LELARASMEE, A. RUEHLI, AND A. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. Compt.-Aided Design Integr. Circuits Syst., 1 (1982), pp. 131–145.
[31] E. LINDELÖF, *Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires*, Journal de Mathématiques Pures et Appliquées, (1894), pp. 117–128.
[32] J.-L. LIONS, Y. MADAY, AND G. TURINICI, *A parareal in time discretization of PDEs*, C.R. Acad. Sci. Paris, Série I, 332 (2001), pp. 661–668.
[33] P.-L. LIONS, *On the Schwarz alternating method I*, in First International Symposium on Domain Decomposition Methods for PDEs, Philadelphia, 1988, pp. 1–42.
[34] B. C. MANDAL, *A Time-Dependent Dirichlet-Neumann Method for the Heat Equation*, in Domain Decomposition in Science and Engineering XXI, J. Erhel, M. J. Gander, L. Halpern, G. Pichot, T. Sassi, and O. B. Widlund, eds., vol. 98, Springer-Verlag, 2014, pp. 467–475.
[35] L. MARTINI AND A. QUARTERONI, *An Iterative Procedure for Domain Decomposition Methods: a Finite Element Approach*, in SIAM, in Domain Decomposition Methods for PDEs, I, R. Glowinski et al., ed., Philadelphia, 1988, pp. 129–143.
[36] ———, *A Relaxation Procedure for Domain Decomposition Method using Finite Elements*, Numer. Math., 55 (1989), pp. 575–598.
[37] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problems*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. 459–482.
[38] O. NEVANLINNA, *Remarks on Picard-Lindelöf iterations part I*, BIT, 29 (1989), pp. 328–346.
[39] ———, *Remarks on Picard-Lindelöf iterations part II*, BIT, 29 (1989), pp. 535–562.
[40] F. OBERHETTINGER AND L. BADII, *Tables of Laplace Transforms*, Springer-Verlag, 1973.
[41] B. ONG, S. HIGH, AND F. KWOK, *Pipeline schwarz waveform relaxation*, in Domain Decomposition Methods in Computational Science and Engineering XXII, Springer-Verlag, 2016, pp. 363–370.
[42] E. PICARD, *Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires*, Journal de Mathématiques Pures et Appliquées, (1893), pp. 217–272.
[43] A. QUARTERONI AND A. VALLI, *Domain Decomposition Methods for Partial Differential Equations*, Clarendon Press, 1999.
[44] F. SHAKIB AND T. J. R. HUGHES, *A new finite element formulation for computational fluid dynamics: IX. Fourier analysis of space-time Galerkin/least-squares algorithms*, Comput. Methods Appl. Mech. Eng., 87 (1991), pp. 35–58.

[45] J. J. SUDIRHAM, J. J. W. VAN DER VEGT, AND R. M. J. VAN DAMME, *Space-time discontinuous Galerkin method for advection-diffusion problems on time-dependent domains*, Appl. Numer. Math., 56 (2006), pp. 1491–1518.
[46] A. TOSELLI AND O. B. WIDLUND, *Domain Decomposition Methods, Algorithms and Theory*, Springer, 2005.