SIAM J. NUMER. ANAL.
Vol. 0, No. 0, pp. 000–000

OPTIMIZED SCHWARZ WAVEFORM RELAXATION METHODS FOR ADVECTION REACTION DIFFUSION PROBLEMS*

M. J. GANDER† AND L. HALPERN‡

Abstract. We study in this paper a new class of waveform relaxation algorithms for large systems of ordinary differential equations arising from discretizations of partial differential equations of advection reaction diffusion type. We show that the transmission conditions between the subsystems have a tremendous influence on the convergence speed of the waveform relaxation algorithms, and we identify transmission conditions with optimal performance. Since these optimal transmission conditions are expensive to use, we introduce a class of local transmission conditions of Robin type, which approximate the optimal ones and can be used at the same cost as the classical transmission conditions. We determine the transmission conditions in this class with the best performance of the associated waveform relaxation algorithm. We show that the new algorithm is well posed and converges much faster than the classical one. We illustrate our analysis with numerical experiments.

Key words. domain decomposition, waveform relaxation, Schwarz methods, time parallelism

AMS subject classifications. AU MUST PROVIDE

DOI. 10.1137/050642137

1. Introduction.

Waveform relaxation algorithms have been invented to solve extremely large systems of ordinary differential equations arising in circuit simulation [26]. They use a partition of the original problem into subproblems, which are then solved separately, and an iteration with information exchange between subproblems leads to the solution of the original problem. Since the solution of the subproblems can be done in parallel, these algorithms are very well suited for implementations on parallel computers. The main drawback of waveform relaxation algorithms is in general their slow convergence, for a review, see [1].

There are two main classical approaches in the literature to solve parabolic problems in parallel. The first approach consists of discretizing the equations uniformly in time with an implicit scheme and then applying a domain decomposition technique to the elliptic problems obtained at each time step, see, for example, [2, 32, 3] and references therein. This approach has the disadvantage that a uniform time discretization needs to be enforced across different subdomains, and one thus loses one of the main features of domain decomposition algorithms, namely, to be able to treat different subdomains numerically differently with an adapted procedure for each subdomain. A second disadvantage is that small amounts of information need to be exchanged at every time level, which can be costly in a parallel environment where the startup cost to send information is significant. In addition, the iteration in time cannot proceed before all the subdomains have converged. An interesting variant, which avoids iterating by making explicit predictions at the interfaces, can be found in [36].

The second classical approach consists of discretizing the equations in space first and then applying a waveform relaxation algorithm to the large system of ordinary

*Received by the editors October 7, 2005; accepted for publication (in revised form) October 6, 2006; published electronically DATE.
http://www.siam.org/journals/sinum/x-x/64213.html

†Section de Mathématiques, Université de Genève, 2-4 rue du Lièvre, CP 64, CH-1211 Genève, Switzerland (martin.gander@unige.ch).
‡LAGA, Institut Galilée, Université Paris XIII, 93430 Villetaneuse, France (halpern@math.univ-paris13.fr).

differential equations obtained from the spatial discretization. A formulation using discretized subdomains can be found in [25]. The main disadvantage of this approach is that spatial information of the connectivity of the subsystems in the large system of ordinary differential equations is lost, and parameters like physical overlap and information exchange are difficult to interpret in this context. Using a different approach and abandoning the idea of subsystems, efficient waveform relaxation algorithms of multigrid type, see [29, 38, 21, 22], and also convolution waveform relaxation algorithms, see [20, 23], have been introduced and analyzed.

To avoid the inherent problems of the classical decomposition approaches, waveform relaxation algorithms for problems in space-time were formulated in [16, 14, 13] and independently in [18] directly at the continous level without discretization. The spatial domain is decomposed into subdomains, and time dependent problems are solved iteratively on subdomains, exchanging information at the interfaces between subdomains. This approach permits a different numerical treatment in both space and time of the subdomain problems, and information is exchanged less often between subdomains. The iteration is defined as in the classical Schwarz case, but as in waveform relaxation, time dependent subproblems are solved, which explains the names of these methods. Unfortunately these algorithms, although robust with respect to refinement, if the overlap is held constant, are still converging only slowly.

We show in this paper for a model problem of advection reaction diffusion type why the convergence of the Schwarz waveform relaxation algorithm is slow. By analyzing the convergence behavior of the classical overlapping Schwarz waveform relaxation algorithm applied to the model problem, we show that the classical algorithm is using ineffective transmission conditions. The classical transmission conditions inhibit the information exchange between subdomains and therefore slow down the convergence of the algorithm. Using ideas introduced in [10], we derive optimal transmission conditions for the Schwarz waveform relaxation algorithm. These transmission conditions coincide with the transparent boundary conditions used to truncate computational domains, which were first studied in [7] for hyperbolic problems and in [19] for advection diffusion problems. Transparent transmission conditions lead to Schwarz waveform relaxation algorithms which converge in a finite number of steps, equal to the number of subdomains, see [11] for the wave equation case. In general, however, the transparent boundary conditions are expensive to compute since they involve nonlocal operators. Similar to the approach for stationary problems in [34, 24, 8, 17], we approximate the transparent transmission conditions locally at the interfaces between subdomains, see [10, 12, 31]. This leads to algorithms which converge even without overlap and are well suited to couple different numerical methods, like finite element and finite differences methods, see [4]. We then optimize the convergence rate, including an overlap in the optimization if desired.

Since the algorithms are eventually discretized to be used on a parallel computer, we analyze the performance of the optimized algorithms asymptotically as the discretization parameter goes to zero. This analysis reveals an interesting relationship between the space-time discretization (implicit-explicit) and the convergence of the optimized algorithm. Numerical experiments show that the convergence rates are improved by orders of magnitude over the rate of the classical overlapping Schwarz waveform relaxation methods.

This paper is organized as follows: In section 2, we present the model problem for which we study the overlapping Schwarz waveform relaxation algorithm in what follows. We include fundamental existence results for the solution, which are later used to prove well posedness and convergence of the algorithm. In section 3, we introduce the classical overlapping Schwarz waveform relaxation algorithm, show that it is well posed when applied to the advection reaction diffusion equation, and analyze its convergence. In section 4, we introduce the optimal Schwarz waveform relaxation algorithm, an algorithm that uses, instead of Dirichlet transmission conditions, transparent ones. Because such transmission conditions can be expensive to use, we also introduce local approximations of these transmission conditions. The core of this paper is contained in section 5, where we analyze the optimized Schwarz waveform relaxation algorithm with Robin transmission conditions. We show that the algorithm is well posed and convergent. We also derive the optimal parameters in the Robin transmission conditions and their dependence on the problem parameters, and we study the asymptotic dependence of the discretized algorithm on the mesh parameters. In section 6, we show numerical results for the classical and optimized Schwarz waveform relaxation algorithms, which show how drastically the convergence behavior is improved using optimized transmission conditions. We present our conclusions in section 7. All our analysis is performed for the simple case of a two subdomain decomposition, since we improve the algorithm locally between subdomains. We show, however, numerical experiments for more than two subdomains, which indicate that the results of our analysis are valid in that case as well.

2. Model problem and function spaces. Our guiding example is the one dimensional advection reaction diffusion equation

$$\mathcal{L}u := \partial_t u - \nu\partial_{xx}u + a\partial_x u + bu = f \quad \text{in } \Omega \times (0,T), \tag{2.1}$$

where $\Omega = \mathbb{R}$, $\nu > 0$, and a and b are constants which do not both vanish simultaneously. The case of the heat equation needs special treatment and can be found in [12]. Without loss of generality, we can assume that the advection coefficient a is nonnegative since $a < 0$ amounts to changing x into $-x$. We can also assume that the reaction coefficient b is nonnegative. If not, a change of variables $v = ue^{-\sigma t}$ with $\sigma + b > 0$ will lead to (2.1) with a positive reaction coefficient.

A weak solution of (2.1) for $\Omega = \mathbb{R}$ is defined to be a $u \in \mathcal{C}(0,T;L^2(\Omega)) \cap L^2(0,T;H^1(\Omega))$ such that for any v in $H^1(\Omega)$

$$\frac{d}{dt}(u,v) + \nu(\partial_x u, \partial_x v) + \frac{a}{2}((\partial_x u, v) - (\partial_x v, u)) + b(u,v) = (f,v) \text{ in } \mathcal{D}'(0,T), \tag{2.2}$$

where $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$. Problem (2.1) is completed by the initial condition

$$u(x,0) = u_0(x) \quad \text{in } \Omega. \tag{2.3}$$

We have a first existence and uniqueness result, proved in [27].

THEOREM 2.1. *For $\Omega = \mathbb{R}$, if the initial value u_0 is in $L^2(\Omega)$ and the right-hand side f is in $L^2(0,T;L^2(\Omega))$, then there exists a unique weak solution u of* (2.1), (2.3).

With the transmission conditions we will introduce later, we will need more regularity in our analysis, in the anisotropic Sobolev spaces $H^{r,s}(\Omega \times (0,T)) = L^2(0,T;H^r(\Omega)) \cap H^s(0,T;L^2(\Omega))$ defined in [27].

THEOREM 2.2. *For $\Omega = \mathbb{R}$, if the initial value u_0 is in $H^1(\Omega)$ and the right-hand side f is in $L^2(0,T;L^2(\Omega))$, then the weak solution u of* (2.1), (2.3) *is in $H^{2,1}(\Omega \times (0,T))$. If u_0 is in $H^2(\Omega)$ and f is in $H^{1,1/2}(\Omega \times (0,T))$, then u is in $H^{3,3/2}(\Omega \times (0,T))$.*

For the proof, and the trace theorems in $H^{r,s}$, we refer to [27].

3. Classical Schwarz waveform relaxation. We decompose the spatial domain $\Omega = \mathbb{R}$ into two overlapping subdomains $\Omega_1 = (-\infty, L)$ and $\Omega_2 = (0, \infty)$, $L > 0$. The overlapping Schwarz waveform relaxation algorithm consists then of solving iteratively subproblems on $\Omega_1 \times (0,T)$ and $\Omega_2 \times (0,T)$ using as a boundary condition at the interfaces $x = 0$ and $x = L$ the values obtained from the previous iteration. The algorithm is thus for iteration index $k = 1, 2, \dots,$ given by

$$
\begin{array}{rcll rcll}
\mathcal{L}u_1^k &=& f & \text{in } \Omega_1 \times (0,T), & \mathcal{L}u_2^k &=& f & \text{in } \Omega_2 \times (0,T),\\
u_1^k(\cdot,0) &=& u_0 & \text{in } \Omega_1, & u_2^k(\cdot,0) &=& u_0 & \text{in } \Omega_2,\\
u_1^k(L,\cdot) &=& u_2^{k-1}(L,\cdot) & \text{in } (0,T), & u_2^k(0,\cdot) &=& u_1^{k-1}(0,\cdot) & \text{in } (0,T),
\end{array}
\tag{3.1}
$$

where an initial guess $u_1^0(0,t)$ and $u_2^0(L,t)$, $t \in (0,T)$, needs to be provided. We first study the well posedness of algorithm (3.1) and then analyze its convergence properties. While algorithm (3.1) is also well defined without overlap, $L = 0$, it is not convergent, since no information is exchanged in that case. This will be different for the optimized algorithms proposed in section 5.

3.1. Well posedness of the algorithm. We first need to show the well posedness of each subdomain problem. Without loss of generality, we consider the subdomain problem on Ω_1 only:

$$
\begin{array}{rcll}
\mathcal{L}v &=& f & \text{in } \Omega_1 \times (0,T),\\
v(\cdot,0) &=& u_0 & \text{in } \Omega_1,\\
v(L,\cdot) &=& g & \text{in } (0,T).
\end{array}
\tag{3.2}
$$

THEOREM 3.1. *If $f \in L^2(0,T;L^2(\Omega_1))$, $u_0 \in H^1(\Omega_1)$, and $g \in H^{\frac{3}{4}}(0,T)$ and if the compatibility condition*

$$u_0(L) = g(0) \tag{3.3}$$

is satisfied, then problem (3.2) *has a unique solution v in $H^{2,1}(\Omega_1 \times (0,T))$. Moreover $v(0,\cdot)$ is in $H^{\frac{3}{4}}(0,T)$, and the following compatibility property holds:*

$$\lim_{t \to 0+} v(0,t) = u_0(0). \tag{3.4}$$

Proof. The proof of existence and uniqueness in $H^{2,1}(\Omega_1 \times 0,T))$ can be found in [27]. The compatibility relation follows from the trace theorem in [27]. □

By the Sobolev embedding theorem, u_0 is continuous on $\bar{\Omega}_1$ and g is continuous on $[0,T]$, which gives a classical meaning to the compatibility condition (3.3). The preceding result ensures that the subdomain problems are well posed in the classical algorithm, provided the initial and boundary conditions satisfy the compatibility condition (3.3) for each iteration step.

To show that this is indeed the case, let $u_2^0(L,\cdot)$ and $u_1^0(0,\cdot)$ be given in $H^{\frac{3}{4}}(0,T)$ such that $u_2^0(L,\cdot) = u_0(L)$ and $u_1^0(0,\cdot) = u_0(0)$. Then, by Theorem 3.1, the first iteration of the overlapping Schwarz waveform relaxation algorithm (3.1) defines a unique first iterate (u_1^1, u_2^1) in $H^{2,1}(\Omega_1 \times (0,T)) \times H^{2,1}(\Omega_2 \times (0,T))$. Furthermore, $u_1^1(0,\cdot)$ and $u_2^1(L,\cdot)$ are in $H^{\frac{3}{4}}(0,T)$, $\lim_{t\to 0+} u_1^1(0,t) = u_0(0)$, and $\lim_{t \to 0+} u_2^1(L,t) = u_0(L)$. Hence by induction, the algorithm is well posed.

3.2. Convergence of the algorithm. By linearity, the error between the solution u and the iterates u_j^k, $j = 1, 2$, of the overlapping Schwarz waveform relaxation algorithm (3.1) satisfy a homogeneous advection reaction diffusion equation with a homogeneous initial condition. We therefore study in what follows the homogeneous problem with data on the interfaces only. Let h_L and h_0 be given in $H^{\frac{3}{4}}(0,T)$ with $h_L(0)=0$ and $h_0(0)=0$, to satisfy the compatibility conditions, and let (e_1,e_2) be the solution in $H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))$ of the equations

$$\begin{array}{rclll} \mathcal{L}e_1 &=& 0 & \text{in } \Omega_1\times(0,T), \\ e_1(\cdot,0) &=& 0 & \text{in } \Omega_1, \\ e_1(L,\cdot) &=& h_L & \text{in } (0,T), \end{array} \qquad \begin{array}{rclll} \mathcal{L}e_2 &=& 0 & \text{in } \Omega_2\times(0,T), \\ e_2(\cdot,0) &=& 0 & \text{in } \Omega_2, \\ e_2(0,\cdot) &=& h_0 & \text{in } (0,T). \end{array} \tag{3.5}$$

Our analysis is based on the Fourier transform, which we denote for any function $h\in L^2(\mathbb{R})$ by $\hat{h}:=\mathcal{F}h$. We define the one-sided space ${}_0H^{\frac{3}{4}}(0,T)=\{\varphi\in H^{\frac{3}{4}}(0,T),\varphi(0)=0\}$, equipped with the norm $\|\varphi\|_{{}_0H^{\frac{3}{4}}(0,T)}:=\inf\ \{\ \|\Phi\|_{H^{\frac{3}{4}}(\mathbb{R})},\Phi=\varphi$ a.e. in $(0,T),\Phi=0$ a.e. in $(-\infty,0)\}$.

LEMMA 3.2. *Let $L>0$. If $a>0$ or $a=0$ and $b>0$, then the map $\mathcal{G}_D$ associated with equations* (3.5)*,*

$$\mathcal{G}_D:(h_L,h_0)\mapsto(e_2(L,\cdot),e_1(0,\cdot)), \tag{3.6}$$

is defined from $({}_0H^{\frac{3}{4}}(0,T))^2$ into itself, and $\mathcal{G}_D^2$ is a strict contraction on $({}_0H^{\frac{3}{4}}(0,T))^2$.

Proof. Since h_L and h_0 are in ${}_0H^{\frac{3}{4}}(0,T)$, we can extend them in $H^{\frac{3}{4}}(\mathbb{R})$ to obtain $\tilde{h}_L$ and $\tilde{h}_0$ vanishing on $(-\infty,0)$. We then extend equations (3.5) in time to $\mathbb{R}$, and their solution coincides with (e_1,e_2) on $(0,T)$. Therefore, we still call it (e_1,e_2). By Fourier transform in time, we find in each subdomain the same ordinary differential equation

$$i\omega\hat{e}_j-\nu\partial_{xx}\hat{e}_j+a\partial_x\hat{e}_j+b\hat{e}_j=0,\quad j=1,2, \tag{3.7}$$

with the characteristic roots

$$r^+=\frac{a+\sqrt{d}}{2\nu},\quad r^-=\frac{a-\sqrt{d}}{2\nu},\quad d=a^2+4\nu(b+i\omega), \tag{3.8}$$

where $\sqrt{d}$ is the complex square root with positive real part. Therefore, $\Re(r^+)>0$ and $\Re(r^-)<0$, and we find, using that e_j is in $L^2(\Omega_j)$,

$$\hat{e}_1(x,\omega)=\mathcal{F}\tilde{h}_L(\omega)e^{r^+(x-L)},\qquad \hat{e}_2(x,\omega)=\mathcal{F}\tilde{h}_0(\omega)e^{r^-x}. \tag{3.9}$$

On the interfaces of the subdomains, we therefore have

$$\mathcal{F}(\mathcal{G}_D(\tilde{h}_L,\tilde{h}_0))(\omega)=(\mathcal{F}\tilde{h}_0(\omega)e^{r^-L},\mathcal{F}\tilde{h}_L(\omega)e^{-r^+L}).$$

Since $\tilde{h}_0$ and $\tilde{h}_L$ vanish in $\mathbb{R}_-$, their Fourier transforms are analytic in the half-plane $\Im\tau<0$, and by (3.9) and the Paley–Wiener theorem [37], $e_1(0,\cdot)$ and $e_2(L,\cdot)$ vanish in $\mathbb{R}_-$. Since they are in $H^{\frac{3}{4}}(\mathbb{R})$, they are continuous, and therefore $e_2(L,0)=0$ and $e_1(0,0)=0$: the map $\mathcal{G}_D$ maps $({}_0H^{\frac{3}{4}}(0,T))^2$ into itself. We have furthermore

$$\mathcal{F}(\mathcal{G}_D^2(\tilde{h}_L,\tilde{h}_0))(\omega)=e^{(r^--r^+)L}(\mathcal{F}\tilde{h}_0(\omega),\mathcal{F}\tilde{h}_L(\omega)). \tag{3.10}$$

Denoting by

$$C_D:=\sup_{\omega\in\mathbb{R}}e^{(r^--r^+)L}=e^{-\frac{L}{\nu}(\sqrt{a^2+4\nu b})}, \tag{3.11}$$

we get for any extension $(\tilde{h}_0, \tilde{h}_L)$ of (h_0, h_L)

$$\|\mathcal{G}_D^2(h_L, h_0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2} \le \|\mathcal{G}_D^2(\tilde{h}_L, \tilde{h}_0)\|_{(H^{\frac{3}{4}}(\mathbb{R}))^2} \le C_D\|(\tilde{h}_L, \tilde{h}_0)\|_{(H^{\frac{3}{4}}(\mathbb{R}))^2}.$$

Taking the infimum on all the extensions on the right-hand side, we get

$$\|\mathcal{G}_D^2(h_L, h_0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2} \le C_D\|(h_L, h_0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2},$$

and since C_D is positive and strictly less than 1, we have proved that $\mathcal{G}_D^2$ is a contraction. □

We now prove convergence of the overlapping Schwarz waveform relaxation algorithm.

THEOREM 3.3. *Let $L > 0$. For $a > 0$ or $a = 0$ and $b > 0$, the iterates (u_1^k, u_2^k) of algorithm* (3.1) *converge to the solution of* (2.1), (2.3) *for any initial guess g_0 and g_L in $H^{\frac{3}{4}}(0,T)$ such that $g_0(0) = u_0(0)$ and $g_L(0) = u_0(L)$.*

Proof. The errors $e_j^k = u_j^k - u$, $j = 1, 2$, satisfy for $k \ge 1$ (3.1) with $f = 0$ and $u_0 = 0$. For positive k, we introduce the interface functions $h_L^k = e_2^k(L, \cdot)$ and $h_0^k = e_1^k(0, \cdot)$ and denote by $h_0^0 = h_0$ and $h_L^0 = h_L$. Using the map $\mathcal{G}_D$, we obtain by induction

$$(h_L^{2k}, h_0^{2k}) = \mathcal{G}_D^{2k}(h_L^0, h_0^0),$$

and thus by Lemma 3.2

$$\|(h_L^{2k}, h_0^{2k})\|_{({}_0H^{\frac{3}{4}}(0,T))^2} \le C_D^k\|(h_L^0, h_0^0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2},$$

with C_D given in (3.11). Solving (3.5) and using (3.9), we obtain for e_1

$$\|e_1\|^2_{L^2(0,T;H^2(\Omega_1))} \le \int_{-\infty}^{\infty} |r^+|^4 \int_{\Omega_1} e^{2\Re(r^+)(x-L)} |\mathcal{F}\tilde{h}_L|^2 dx\, d\omega = \int_{-\infty}^{\infty} \frac{|r^+|^4}{2\Re(r^+)} |\mathcal{F}\tilde{h}_L|^2 d\omega.$$

For $a > 0$ or $a = 0$ and $b > 0$, the denominator in the factor in front of $|\mathcal{F}\tilde{h}_L|^2$ is bounded from below, and for $|\omega|$ large, the factor behaves like $|\omega|^{3/2}$. Therefore

$$\|e_1\|_{L^2(0,T;H^2(\Omega_1))} \le C\|\tilde{h}_L\|_{H^{\frac{3}{4}}(\mathbb{R})},$$

and the same result also holds for $\|e_1\|_{H^1(0,T;L^2(\Omega_1))}$. Hence

$$\|e_1\|_{H^{2,1}((\Omega_1)\times(0,T))} \le M_D\|h_L\|_{{}_0H^{\frac{3}{4}}(0,T)}, \tag{3.12}$$

and similarly for e_2. Now we apply (3.12) to the errors e_1^{2k+1} and e_2^{2k+1} in the iteration and obtain

$$\begin{aligned}&\|(e_1^{2k+1}, e_2^{2k+1})\|_{H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))}\\ &\qquad \le M_D\|(h_L^{2k}, h_0^{2k})\|_{({}_0H^{\frac{3}{4}}(0,T))^2} \le M_D C_D^k\|(g_L - u(L,\cdot), g_0 - u(0,\cdot))\|_{({}_0H^{\frac{3}{4}}(0,T))^2},\end{aligned}$$

which together with Lemma 3.2 completes the proof. A similar argument also holds for even iteration numbers. □

Theorem 3.3 shows that the overlapping Schwarz waveform relaxation algorithm converges and that the convergence rate is at least linear and is independent of the length of the time interval. It does however depend on the problem parameters ν,

a, and b and the overlap L. Using also the preceding Lemma 3.2, the error in the overlapping Schwarz waveform relaxation algorithm satisfies on the interfaces over a double iteration step in Fourier space relation (3.10) or equivalently

$$\hat{e}_1^{k+1}(L,\omega) = \rho_D \hat{e}_1^{k-1}(L,\omega), \quad \hat{e}_2^{k+1}(0,\omega) = \rho_D \hat{e}_2^{k-1}(0,\omega), \tag{3.13}$$

where the convergence factor $\rho_D = \rho_D(\omega, L, \nu, a, b)$ is given by

$$\rho_D(\omega, L, \nu, a, b) := e^{(r^- - r^+)L} = e^{-\frac{\sqrt{a^2+4\nu(b+i\omega)}}{\nu}L}. \tag{3.14}$$

Note that the convergence factor ρ_D is uniformly bounded in modulus for all ω by a quantity strictly less than 1,

(3.15)

$$R_D(\omega, L, \nu, a, b) := |\rho_D(\omega, L, \nu, a, b)| \leq \bar{R}_D(L, \nu, a, b) := R_D(0, L, \nu, a, b) = e^{-\frac{\sqrt{a^2+4\nu b}}{\nu}L},$$

and for L small, we have

$$\bar{R}_D = 1 - \frac{\sqrt{a^2+4\nu b}}{\nu} L + O(L^2). \tag{3.16}$$

Using the convergence factor ρ_D from Fourier analysis allows us to obtain a sharper convergence result for bounded time intervals.

THEOREM 3.4 (superlinear convergence). *For the advection reaction diffusion equation on a bounded time interval* $(0, T)$*, the asymptotic convergence rate of the overlapping Schwarz waveform relaxation algorithm* (3.1) *is superlinear:*

$$||e_j^{2k}(0,\cdot)||_{L^\infty(0,T)} \leq erfc\left(\frac{kL}{\sqrt{\nu T}}\right) ||e_j^0(0,\cdot)||_{L^\infty(0,T)}, \quad j = 1, 2,$$

where the error function complement is defined by $erfc(x) := \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-s^2} ds$.

Proof. By induction on the relations (3.13), we obtain

$$\hat{e}_1^{2k}(0,\omega) = \rho_D^k \hat{e}_1^0(0,\omega), \qquad \hat{e}_2^{2k}(L,\omega) = \rho_D^k \hat{e}_2^0(L,\omega). \tag{3.17}$$

Using the inverse Fourier transform and the convolution theorem, we find

$$e_1^{2k}(0,t) = (\mathcal{F}^{-1}\rho_D^k) * e_1^0(0,t), \qquad e_2^{2k}(L,t) = (\mathcal{F}^{-1}\rho_D^k) * e_2^0(L,t). \tag{3.18}$$

Now the inverse Fourier transform of ρ_D^k is

$$\mathcal{F}^{-1}\rho_D^k = \frac{kL}{\sqrt{\nu\pi} t^{3/2}} e^{-\frac{(kL)^2}{\nu t} - \left(\frac{a^2}{4\nu} + b\right)t},$$

and we can therefore estimate for $j = 1, 2$

$$\|e_j^{2k}(0,\cdot)\|_{L^\infty(0,T)} \leq \|\mathcal{F}^{-1}\rho_D^k\|_{L^1(0,T)} \|e_j^0(0,\cdot)\|_{L^\infty(0,T)} \leq erfc\left(\frac{kL}{\sqrt{\nu T}}\right) \|e_j^0(0,\cdot)\|_{L^\infty(0,T)},$$

where the last inequality follows from estimating the term $e^{-(\frac{a^2}{4\nu}+b)t}$ by 1. By a similar argument for the second subdomain, the result follows. ∎

This result was first proved for bounded domains in [18], see also [15]. It also holds in higher dimensions and for general decompositions, for the heat equation, see [14],

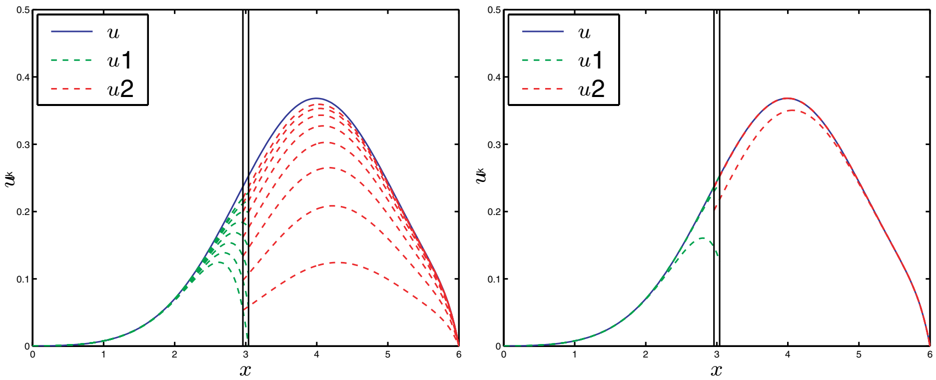

FIG. 3.1. *On the left, the first few iterates of the classical Schwarz waveform relaxation algorithm (dashed) at the end of the time interval $t = T$, together with the exact solution (solid), and on the right the first iterates of the new optimized Schwarz waveform relaxation algorithm.*

and for advection diffusion, see [5]. The result differs significantly from the classical linear convergence result of the overlapping Schwarz method for elliptic problems and also from the classical superlinear convergence results for waveform relaxation, which is slower, see [35]. Furthermore, one can show that the convergence rate depends only on the number of subdomains in higher order terms, see [15], and hence coarse grid preconditioners are not necessary for evolution problems of this type.

The Dirichlet transmission conditions at the interfaces are however responsible for slow convergence in the classical Schwarz waveform relaxation algorithm: in Figure 3.1 on the left, we show the first few iterations at the end of the time interval for a model problem. On the right, we show the first few iterations of the new, much faster algorithm we will develop in what follows.

4. Optimal Schwarz waveform relaxation. We now introduce transmission conditions which are more effective for the information exchange between subdomains. The new algorithm is

(4.1)
$$\begin{array}{rlrl} \mathcal{L}u_1^k &= f \quad \text{in } \Omega_1 \times (0,T), & \mathcal{L}u_2^k &= f \quad \text{in } \Omega_2 \times (0,T),\\ u_1^k(\cdot,0) &= u_0, & u_2^k(\cdot,0) &= u_0,\\ (\partial_x+\mathcal{S}_1)u_1^k(L,\cdot) &= (\partial_x+\mathcal{S}_1)u_2^{k-1}(L,\cdot), & (\partial_x+\mathcal{S}_2)u_2^k(0,\cdot) &= (\partial_x+\mathcal{S}_2)u_1^{k-1}(0,\cdot), \end{array}$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are linear operators in time, possibly pseudodifferential.

4.1. Optimal transmission conditions. The following theorem gives the optimal choice for $\mathcal{S}_1$ and $\mathcal{S}_2$.

THEOREM 4.1. *For $a > 0$ or $a = 0$ and $b > 0$, algorithm* (4.1) *converges to the solution u of* (2.1) *in two iterations for all initial guesses $u_1^0 \in H^{2,1}(\Omega_1 \times (0,T))$ and $u_2^0 \in H^{2,1}(\Omega_2 \times (0,T))$, independently of the size of the overlap $L \geq 0$, if and only if the operators $\mathcal{S}_1$ and $\mathcal{S}_2$ have the corresponding symbols*

$$\sigma_1 = -r^-, \quad \sigma_2 = -r^+, \tag{4.2}$$

where $r^\pm$ are defined in (3.8).

Proof. Using the Fourier transform with parameter ω as in Lemma 3.2, we find for the error

$$\hat{e}_1^k(x,\omega) = \alpha^k(\omega)e^{r^+(x-L)}, \quad \hat{e}_2^k(x,\omega) = \beta^k(\omega)e^{r^-x}, \quad k \geq 1, \tag{4.3}$$

where α^k and β^k are constants, which, using the new transmission conditions, satisfy for $k \geq 1$ the recurrence relation

$$\begin{aligned} \alpha^{k+1}(r^+ + \sigma_1) &= \beta^k(r^- + \sigma_1)e^{r^- L}, \\ \beta^{k+1}(r^- + \sigma_2) &= \alpha^k(r^+ + \sigma_2)e^{-r^+ L}. \end{aligned} \tag{4.4}$$

Now for an arbitrary initial guess u_1^0 and u_2^0, the coefficients α^1 and β^1 will in general not vanish. Since $r^- + \sigma_1 = r^+ + \sigma_2 = 0$ implies $r^+ + \sigma_1 \neq 0$ and $r^- + \sigma_2 \neq 0$, we obtain from (4.4) that α^2 and β^2 are identically zero if and only if $r^- + \sigma_1 = r^+ + \sigma_2 = 0$. ∎

Note that the symbols σ_1, σ_2 given in (4.2) are not polynomials in $i\omega$, and hence the optimal corresponding transmission operators $\mathcal{S}_1$, $\mathcal{S}_2$ are nonlocal operators in time; they correspond to integral transfer operators in time along the interfaces between subdomains. Even though such operators can be efficiently implemented, see, for example, [30], they are more costly than local transfer operators and the latter are in general preferred. It is therefore of interest to approximate the nonlocal operators by local ones, whose symbols are polynomials in $i\omega$. Using each equation in (4.4) at iteration k in the other one at iteration $k+1$, we find

$$\alpha^{k+1} = \rho\alpha^{k-1}, \quad \beta^{k+1} = \rho\beta^{k-1}$$

with the new convergence factor

$$\rho = \frac{r^- + \sigma_1}{r^+ + \sigma_1} \cdot \frac{r^+ + \sigma_2}{r^- + \sigma_2} e^{(r^- - r^+)L}. \tag{4.5}$$

4.2. Approximations of the optimal transmission conditions. We approximate the symbols σ_1 and σ_2 from (4.2) corresponding to the optimal transmission operators by constants, which leads to Robin transmission conditions in the Schwarz waveform relaxation algorithm (4.1), i.e.,

$$\mathcal{S}_1 := \frac{-a + p}{2\nu}, \qquad \mathcal{S}_2 := \frac{-a - p}{2\nu}. \tag{4.6}$$

The choice of the parameter p is restricted by the requirement that the subdomain problems need to be well posed, and a good choice should lead to a rapidly converging algorithm; both issues we will analyze in detail in the following section.

Notice that using the knowledge of the symbols (4.2) of the optimal transmission conditions, we have chosen a particular form for the low order approximation, leading to (4.6). In general one is not required to do so; in particular, we could have chosen, for example, two different parameters p in (4.6), which would have given more freedom in the optimization process we study later, or even choosen a different p at each iteration, as it was done for a steady problem in [9]. One could also choose higher order transmission conditions, i.e., approximations by polynomials in $i\omega$. Having one parameter only however greatly simplifies the optimization process, so we leave the more general cases for future studies.

5. Optimized Schwarz waveform relaxation. We now study the Schwarz waveform relaxation algorithm with Robin transmission conditions. We start with the overlapping case, $L > 0$. We first show under what conditions on the free parameter p the algorithm is well posed and then prove convergence of the algorithm for a general choice of p satisfying these conditions. We also study the influence of p on the performance of the algorithm and propose two choices for p: one choice motivated by a low frequency approximation and a second choice which optimizes the performance of the algorithm. We then prove that the algorithm converges also without overlap, and we again study the influence of p on the performance of the nonoverlapping algorithm.

5.1. Well posedness of the algorithm. As in the case of the classical Schwarz waveform relaxation algorithm studied in section 3, we first need to analyze under which conditions the subdomain problems of the algorithm with Robin transmission conditions is well posed. Without loss of generality, we study only the well posedness of the subdomain problem on Ω_1:

$$\begin{aligned} \mathcal{L}v &= f && \text{in } \Omega_1 \times (0,T), \\ v(\cdot,0) &= u_0 && \text{in } \Omega_1, \\ (\partial_x v + \mathcal{S}_1 v)(L,\cdot) &= g_L && \text{in } (0,T). \end{aligned} \tag{5.1}$$

We first show an extension result, which allows us to reduce the study of the well posedness to the case with homogeneous initial and boundary conditions.

LEMMA 5.1. *If u_0 is in $H^1(\Omega_1)$ and g_L is in $H^{\frac{1}{4}}(0,T)$, then there exists an extension w in $H^{2,1}(\Omega_1 \times (0,T))$ such that $w(\cdot,0) = u_0$ in Ω_1 and $(\partial_x w + \mathcal{S}_1 w)(L,\cdot) = g_L$ on $(0,T)$.*

Proof. Let $\tilde{g}_L$ be in $H^{\frac{3}{4}}(0,T)$ such that $\tilde{g}_L(0) = u_0(L)$. By the continuous extension theorem, there exists a w_1 in $H^{2,1}(\Omega \times (0,T))$ such that

$$w_1(\cdot,0) = u_0, \quad w_1(L,\cdot) = \tilde{g}_L, \quad \partial_x w_1(L,\cdot) = 0$$

and a w_2 in $H^{2,1}(\Omega \times (0,T))$ such that

$$w_2(\cdot,0) = 0, \quad w_2(L,\cdot) = 0, \quad \partial_x w_2(L,\cdot) = g_L - \mathcal{S}_1 \tilde{g}_L.$$

Now the sum $w := w_1 + w_2$ is the desired extension in $H^{2,1}(\Omega \times (0,T))$. □

Thus it suffices to analyze the well posedness of the problem with homogeneous initial and boundary conditions:

$$\begin{aligned} \mathcal{L}\tilde{v} &= F && \text{in } \Omega_1 \times (0,T), \\ \tilde{v}(\cdot,0) &= 0 && \text{in } \Omega_1, \\ (\partial_x \tilde{v} + \mathcal{S}_1 \tilde{v})(L,\cdot) &= 0 && \text{in } (0,T), \end{aligned} \tag{5.2}$$

where $\tilde{v} = v - w$ and the right-hand side function $F = f - \mathcal{L}w$ is in $L^2(0,T;L^2(\Omega_1))$ if f is in $L^2(0,T;L^2(\Omega_1))$. We start with the weak formulation: for any φ in $H^1(\Omega_1)$, we multiply the equation by φ, integrate, and use Green's formula and the boundary condition to obtain in $\mathcal{D}'(0,T)$

$$\frac{d}{dt}(\tilde{v},\varphi) + \nu(\partial_x \tilde{v}, \partial_x \varphi) + \frac{a}{2}((\partial_x \tilde{v},\varphi) - (\partial_x \varphi, \tilde{v})) + b(\tilde{v},\varphi) + \frac{p}{2}\tilde{v}(L)\varphi(L) = (F,\varphi). \tag{5.3}$$

The following Theorem gives existence, uniqueness, and regularity of the weak solution.

THEOREM 5.2. *Suppose F is in $L^2(0,T;L^2(\Omega_1))$. Then, for any p, problem (5.2) has a unique weak solution $\tilde{v}$ in $H^{2,1}(\Omega_1 \times (0,T))$.*

Proof. The proof is based on a priori estimates.

1. Multiplying equation (5.2) by $\tilde{v}$, integrating in space, and using the boundary condition, we obtain

$$\frac{1}{2}\frac{d}{dt}\|\tilde{v}\|^2 + \nu\|\partial_x \tilde{v}\|^2 + b\|\tilde{v}\|^2 + \frac{p}{2}\tilde{v}^2(L) = (F,\tilde{v}). \tag{5.4}$$

(a) Suppose first $p \geq 0$.
- If $b > 0$, by the Cauchy–Schwarz inequality, and using the inequality

$$\alpha\beta \leq \frac{\eta}{2}\,\alpha^2 + \frac{1}{2\eta}\,\beta^2 \quad \text{for all } \alpha, \beta \in \mathbb{R} \text{ and } \eta > 0 \tag{5.5}$$

in the form $||F||\,||\tilde{v}|| \leq \frac{1}{2b}||F||^2 + \frac{b}{2}||\tilde{v}||^2$, we obtain

$$\frac{1}{2}\frac{d}{dt}\|\tilde{v}\|^2 + \nu\|\partial_x\tilde{v}\|^2 + \frac{b}{2}\|\tilde{v}\|^2 \leq \frac{1}{2b}\|F\|^2,$$

which gives, after integration on any time interval $(0, t)$,

$$\frac{1}{2}||\tilde{v}||^2(t) + \nu\int_0^t ||\partial_x\tilde{v}||^2 + \frac{b}{2}\int_0^t ||\tilde{v}||^2 \leq \frac{1}{2b}\int_0^t ||F||^2. \tag{5.6}$$

- If $b = 0$, we use (5.5) with $\eta = 1$ and get through integration on $(0, t)$

$$\frac{1}{2}\|\tilde{v}\|^2(t) + \nu\int_0^t \|\partial_x\tilde{v}\|^2 \leq \frac{1}{2}\|F\|^2_{L^2(0,T;L^2(\Omega_1))} + \frac{1}{2}\int_0^t \|\tilde{v}\|^2.$$

We then apply the Gronwall lemma and obtain

$$\|\tilde{v}\|^2(t) + 2\nu\int_0^t \|\partial_x\tilde{v}\|^2 \leq e^T\|F\|^2_{L^2(0,T;L^2(\Omega_1))}.$$

(b) Suppose now $p < 0$. We move the boundary term in (5.4) to the right-hand side; using the Sobolev inequality in $H^1(\Omega_1)$,

$$\|\tilde{v}\|^2_{L^\infty(\overline{\Omega_1})} \leq 2\|\partial_x\tilde{v}\|\,\|\tilde{v}\|, \tag{5.7}$$

we bound the boundary term, applying again (5.5),

$$-\frac{p}{2}\tilde{v}^2(L) \leq \frac{\nu}{2}\|\partial_x\tilde{v}\|^2 + \frac{p^2}{2\nu}\|\tilde{v}\|^2;$$

and we conclude using the Gronwall lemma as before.

Thus, in both cases, we have a bound for $\tilde{v}$ in $L^\infty(0,T;L^2(\Omega_1)) \cap L^2(0,T;H^1(\Omega_1))$,

$$\|\tilde{v}\|_{L^\infty(0,T;L^2(\Omega_1))}, \|\tilde{v}\|_{L^2(0,T;H^1(\Omega_1))} \leq C(T)\|F\|_{L^2(0,T;L^2(\Omega_1))}. \tag{5.8}$$

2. To obtain the higher regularity result in the theorem, we need to show that $\partial_x^2\tilde{v}$ and $\partial_t\tilde{v}$ are also in $L^2(0,T;L^2(\Omega_1))$. Multiplying the equation by $-\partial_x^2\tilde{v}$ and integrating in space, we get

$$\frac{1}{2}\frac{d}{dt}||\partial_x\tilde{v}||^2 + \nu||\partial_x^2\tilde{v}||^2 + b||\partial_x\tilde{v}||^2 - \left((\partial_t\tilde{v})\partial_x\tilde{v} + \frac{a}{2}\,(\partial_x\tilde{v})^2 + b(\partial_x\tilde{v})\tilde{v}\right)(L) = -\int_{-\infty}^{L} F\partial_x^2\tilde{v}.$$

Now using the boundary condition to replace $\partial_x\tilde{v}$, we obtain

$$\frac{d}{dt}\left(\frac{1}{2}\|\partial_x\tilde{v}\|^2 + \frac{p-a}{4\nu}\tilde{v}^2(L)\right) + \nu\|\partial_x^2\tilde{v}\|^2 + b\|\partial_x\tilde{v}\|^2 + \frac{p-a}{2\nu}\left(b - \frac{a}{2}\frac{p-a}{2\nu}\right)\tilde{v}^2(L) = -\int_{-\infty}^{L} F\partial_x^2\tilde{v}.$$

Again using the Cauchy–Schwarz inequality and (5.5) on the right, we find after integrating in time

$$(5.9)\qquad \begin{aligned}&\left(\frac{1}{2}\|\partial_x\tilde{v}\|^2+\frac{p-a}{4\nu}\tilde{v}^2(L)\right)(t)+\frac{\nu}{2}\int_0^t\|\partial_x^2\tilde{v}\|^2+b\int_0^t\|\partial_x\tilde{v}\|^2\\&\qquad+\frac{p-a}{2\nu}\left(b-\frac{a}{2}\frac{p-a}{2\nu}\right)\int_0^t\tilde{v}^2(L)\le\frac{1}{2\nu}\int_0^t||F||^2.\end{aligned}$$

First the term $\frac{p-a}{2\nu}\left(b-\frac{a}{2}\frac{p-a}{2\nu}\right)\int_0^t\tilde{v}^2(L)$ is handled as in 1, using (5.7) and (5.8). Then, if $p\ge a$, we obtain

$$\|\partial_x\tilde{v}\|^2+\nu\|\partial_x^2\tilde{v}\|_{L^2(0,T;L^2(\Omega_1))}\le C(T)\|F\|_{L^2(0,T;L^2(\Omega_1))}.$$

If $p<a$, then we pass the term containing $\tilde{v}(L)$ to the right-hand side, and using (5.7), we obtain

$$\frac{1}{2}\|\partial_x\tilde{v}\|^2+\frac{\nu}{2}\int_0^t\|\partial_x^2\tilde{v}\|^2\le\frac{a-p}{4\nu}\left(\alpha||\partial_x\tilde{v}||^2+\frac{1}{\alpha}||\tilde{v}||^2\right)+C(T)\int_0^t||F||^2.$$

Now choosing $\alpha=\nu/(a-p)$ and using (5.8), we obtain

$$\|\partial_x\tilde{v}\|_{L^\infty(0,T;L^2(\Omega_1))},\|\partial_x^2\tilde{v}\|_{L^2(0,T;L^2(\Omega_1))}\le C(T)\|F\|_{L^2(0,T;L^2(\Omega_1))},$$

where we omit the dependence of the constant C on a, p, b, and ν.

Now using (5.2), we have

$$\partial_t\tilde{v}=\nu\partial_x^2\tilde{v}-a\partial_x\tilde{v}-b\tilde{v}+F,$$

and since all the terms on the right-hand side are in $L^2(0,T;L^2(\Omega_1))$ by the previous estimates, it follows that $\partial_t\tilde{v}$ is in $L^2(0,T;L^2(\Omega_1))$, which concludes the a priori estimates in $H^{2,1}(\Omega_1\times(0,T))$. Existence and uniqueness can now be shown using a Galerkin method [27]. ∎

Using Lemma 5.1 and Theorem 5.2, we obtain now the well posedness of the subdomain problems.

THEOREM 5.3. *If f is in $L^2(0,T;L^2(\Omega_1))$, u_0 is in $H^1(\Omega_1)$, and g_L is in $H^{\frac{1}{4}}(0,T)$, then, for any p, problem* (5.1) *has a unique solution v in $H^{2,1}(\Omega_1\times(0,T))$.*

The same result also holds on subdomain Ω_2, the only difference being that $-a$ becomes $+a$ in the estimate (5.9).

THEOREM 5.4. *Let g_L and g_0 be given in $H^{\frac{1}{4}}(0,T)$. If algorithm* (4.1) *with $\mathcal{S}_j$ defined in* (4.6) *is initialized by $(\partial_x u_1^1+\mathcal{S}_1u_1^1)(L,\cdot)=g_L$ and $(\partial_x u_2^1+\mathcal{S}_2u_2^1)(0,\cdot)=g_0$, then, for any p,* (4.1) and (4.6) *define a sequence of iterates (u_1^k,u_2^k) in $H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))$.*

Proof. The proof is done by induction: for $k=1$, (4.1) defines a unique first iterate (u_1^1,u_2^1) in $H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))$ by Theorem 5.3. Assuming now that (u_1^k,u_2^k) is in $H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))$, by the trace theorem, we have that $(\partial_x u_2^k+\mathcal{S}_1u_2^k)(L,\cdot)$ and $(\partial_x u_1^k+\mathcal{S}_2u_1^k)(0,\cdot)$ are in $H^{\frac{1}{4}}(0,T)$, and thus by Theorem 5.3, (u_1^{k+1},u_2^{k+1}) must be in $H^{2,1}(\Omega_1\times(0,T))\times H^{2,1}(\Omega_2\times(0,T))$, which concludes the proof. ∎

For the proof of convergence in the overlapping case, we need however more regularity.

THEOREM 5.5. *For $a > 0$ or $a = 0$ and $b > 0$, let $p \geq 0$ and f be in $H^{1,\frac{1}{2}}(\Omega_1 \times (0,T))$, u_0 be in $H^2(\Omega)$, and g_L be in $H^{\frac{3}{4}}(0,T)$, with the compatibility conditions*

$$g_L(0) = \partial_x u_0(L) + \mathcal{S}_1 u_0(L). \tag{5.10}$$

Then the solution v of the subdomain problem (5.1) is in $H^{3,\frac{3}{2}}(\Omega_1 \times (0,T))$. Furthermore the following compatibility property at $x = 0$ is satisfied:

$$\lim_{t\to 0+}(\partial_x v + \mathcal{S}_2 v)(0,t) = \partial_x u_0(0) + \mathcal{S}_2 u_0(0). \tag{5.11}$$

Proof. In this more regular situation, the solution u of (2.1) is in $H^{3,\frac{3}{2}}(\Omega \times (0,T))$ by Theorem 2.2. Furthermore $\tilde{g}_L = (\partial_x u + \mathcal{S}_1 u)(L,\cdot)$ is in $H^{\frac{3}{4}}(0,T)$. Subtracting u from v, the difference e is in $H^{3,\frac{3}{2}}(\Omega_1 \times (0,T))$, the solution of (5.1) with data $(0,0,h_L = g_L - \tilde{g}_L)$. By Fourier transform, the same calculation as in Lemma 3.6 gives with $\tilde{h}_L$ being an extension of h_L on $\mathbb{R}$ vanishing in $\mathbb{R}_-$

$$\hat{e} = \frac{2\nu}{\sqrt{d}+p}\,\mathcal{F}\tilde{h}_L(\omega)e^{r^+(x-L)}. \tag{5.12}$$

The norm of $\partial_x^3 e$ is therefore given by

$$\|\partial_x^3 e\|^2_{L^2(\Omega_1\times\mathbb{R})} = \int_{\mathbb{R}} \frac{4\nu^2|r^+|^6}{2\Re r^+|\sqrt{d}+p|^2}|\mathcal{F}\tilde{h}_L(\omega)|^2 d\omega,$$

and the norm of e in $H^{\frac{3}{2}}(\mathbb{R}, L^2(\Omega_1))$ is

$$\|e\|^2_{H^{3/2}(\mathbb{R},L^2(\Omega_1))} = \int_{\mathbb{R}} \frac{4\nu^2(1+\omega^2)^{3/2}}{2\Re r^+|\sqrt{d}+p|^2}|\mathcal{F}\tilde{h}_L(\omega)|^2 d\omega.$$

In both cases, for $a > 0$ or $a = 0$ and $b > 0$, and $p \geq 0$, the denominator in the factor in front of $|\mathcal{F}\tilde{h}_L(\omega)|^2$ is bounded from below, and it is easy to see that for large $|\omega|$, it is equivalent to a constant times $|\omega|^{3/2}$. Therefore we have the bound

$$\|e\|_{H^{3,\frac{3}{2}}(\Omega_1\times(0,T))} \leq C\|\tilde{h}_L\|_{H^{\frac{3}{4}}(\mathbb{R})}. \tag{5.13}$$

For the compatibility condition, since $\tilde{h}_L$ vanishes in $\mathbb{R}_-$, its Fourier transform is analytic in the half-plane $\Im\tau < 0$, and by (5.12) and the Paley–Wiener theorem [37], $e(0,\cdot)$ and $\partial_x e(0,\cdot)$ vanish in $\mathbb{R}_-$. Since they are in $H^{\frac{5}{4}}(\mathbb{R})$ and $H^{\frac{3}{4}}(\mathbb{R})$, respectively, they are continuous, and therefore $\lim_{t\to 0+}(\partial_x e + \mathcal{S}_2 e)(0,t) = 0$, which gives the compatibility property for v. ∎

This regularity result shows the well posedness of the algorithm in $H^{3,\frac{3}{2}}(\Omega_1 \times (0,T))$.

THEOREM 5.6. *For $a > 0$ or $a = 0$ and $b > 0$, and $p \geq 0$, let f be in $H^{1,\frac{1}{2}}(\Omega_1 \times (0,T))$, u_0 be in $H^2(\Omega)$, and g_L and g_0 be in $H^{\frac{3}{4}}(0,T)$, with the compatibility conditions*

$$g_L(0) = \partial_x u_0(L) + \mathcal{S}_1 u_0(L), \qquad g_0(L) = \partial_x u_0(0) + \mathcal{S}_2 u_0(0). \tag{5.14}$$

Then, algorithm (4.1) with transmission operators (4.6) defines a sequence of iterates (u_1^k, u_2^k) in $H^{3,\frac{3}{2}}(\Omega_1 \times (0,T)) \times H^{3,\frac{3}{2}}(\Omega_2 \times (0,T))$.

5.2. Convergence of the overlapping algorithm. Let h_L and h_0 be given in ${}_0H^{\frac{3}{4}}(0,T)$. Let (e_1, e_2) be the solution in $H^{3,\frac{3}{2}}(\Omega_1 \times (0,T)) \times H^{3,\frac{3}{2}}(\Omega_2 \times (0,T))$ of the problem

$$\begin{array}{rlrl} \mathcal{L}e_1{=}0 & \text{in } \Omega_1 \times (0,T), & \mathcal{L}e_2{=}0 & \text{in } \Omega_2 \times (0,T), \\ e_1(\cdot,0){=}0 & \text{in } \Omega_1, & e_2(\cdot,0){=}0 & \text{in } \Omega_2, \\ (\partial_x e_1 + \mathcal{S}_1 e_1)(L,\cdot){=}h_L & \text{in } (0,T), & (\partial_x e_2 + \mathcal{S}_2 e_2)(0,\cdot){=}h_0 & \text{in } (0,T). \end{array} \tag{5.15}$$

LEMMA 5.7. *For $a > 0$ or $a = 0$ and $b > 0$, if $p \geq 0$ and $L > 0$, the map $\mathcal{G}_0$ associated with* (5.15)*,*

$$\mathcal{G}_0 : (h_L, h_0) \mapsto ((\partial_x e_2 + \mathcal{S}_1 e_2)(L,\cdot), (\partial_x e_1 + \mathcal{S}_2 e_1)(0,\cdot)), \tag{5.16}$$

is defined from $({}_0H^{\frac{3}{4}}(0,T))^2$ into itself, and $\mathcal{G}_0^2$ is strictly contracting.

Proof. The proof is analogous to the proof of Lemma 3.2 using Fourier analysis. Defining $\tilde{h}_L$ and $\tilde{h}_0$ as any extensions of h_L and h_0 in $H^{\frac{3}{4}}(\mathbb{R})$, vanishing in $\mathbb{R}^-$, we obtain after a short calculation

$$\mathcal{F}(\mathcal{G}_0^2(\tilde{h}_L, \tilde{h}_0))(\omega) = \left(\frac{\sqrt{d} - p}{\sqrt{d} + p}\right)^2 (\mathcal{F}\tilde{h}_0(\omega) e^{(r^- - r^+)L}, \mathcal{F}\tilde{h}_L(\omega) e^{(r^- - r^+)L}),$$

where d and $r^{\pm}$ are defined in (3.8). Since $p \geq 0$, we have $\left|\frac{\sqrt{d}-p}{\sqrt{d}+p}\right| \leq 1$ and thus

$$\|\mathcal{G}_0^2(h_L, h_0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2} \leq C_D \|(h_L, h_0)\|_{({}_0H^{\frac{3}{4}}(0,T))^2},$$

and since C_D defined in (3.11) satisfies $C_D < 1$, the result follows. □

From the proof of this Lemma, we can see that the contraction of the overlapping Schwarz waveform relaxation map with Robin transmission conditions, $\mathcal{G}_0$ given in (5.16), is at least as good as the contraction of the classical map with Dirichlet transmission conditions, $\mathcal{G}_D$ given in (3.6), no matter what one chooses for the parameter $p \geq 0$ in the Robin transmission conditions. Before doing a more thorough comparison, we use the contraction property from Lemma 5.7 to prove convergence of the new algorithm.

THEOREM 5.8. *Let f be in $H^{1,\frac{1}{2}}(\Omega_1 \times (0,T))$ and u_0 be in $H^2(\Omega)$. For $a > 0$ or $a = 0$ and $b > 0$, if $p \geq 0$ and $L > 0$, then the solution (u_1^k, u_2^k) of algorithm* (4.1), (4.6) *converges to the solution u of* (2.1) *for any initial guess $(g_0, g_L) \in (H^{\frac{3}{4}}(0,T))^2$ with the compatibility conditions* (5.14).

Proof. The errors $e_j^k = u_j^k - u$, $j = 1, 2$, satisfy for $k \geq 1$ (4.1) with $f = 0$ and $u_0 = 0$. Introducing the interface functions $h_L^k = (\partial_x e_2^k + \mathcal{S}_1 e_2^k)(L,\cdot)$, $h_0^k = (\partial_x e_1^k + \mathcal{S}_2 e_1^k)(0,\cdot)$ and using the map $\mathcal{G}_0$, we obtain by induction $(h_L^{2k}, h_0^{2k}) = \mathcal{G}_0^{2k}(h_L^0, h_0^0)$, and thus by Lemma 5.7

$$\|(h_L^{2k}, h_0^{2k})\|_{(H^{\frac{3}{4}}(0,T))^2} \leq C_D^k \|(h_L^0, h_0^0)\|_{(H^{\frac{3}{4}}(0,T))^2}.$$

We have by (5.13)

$$\begin{aligned} \|(e_1^{2k+1}, e_2^{2k+1})\|_{H^{3,\frac{3}{2}}(\Omega_1 \times (0,T)) \times H^{3,\frac{3}{2}}(\Omega_2 \times (0,T))} &\leq C \|(h_L^{2k}, h_0^{2k})\|_{(H^{\frac{3}{4}}(0,T))^2} \\ &\leq C C_D^k \|(h_L^0, h_0^0)\|_{(H^{\frac{3}{4}}(0,T))^2}, \end{aligned}$$

which together with Lemma 5.7 completes the proof. □

Having proved convergence, we now compare the performance of the classical Schwarz waveform relaxation algorithm and the new one with Robin transmission conditions. We do this first at the continuous level, which motivates the optimization procedure we introduce in subsections 5.4 and 5.7 for the discretized case. Using Theorem 5.8 and Lemma 5.7, the error in the overlapping Schwarz waveform relaxation algorithm with Robin transmission conditions satisfies on the interfaces over a double iteration step in Fourier the relation

$$\mathcal{F}(\mathcal{G}_0^2(\tilde{h}_L, \tilde{h}_0))(\omega) = \left(\frac{\sqrt{d} - p}{\sqrt{d} + p}\right)^2 e^{(r^- - r^+)L}(\mathcal{F}\tilde{h}_L(\omega), \mathcal{F}\tilde{h}_0(\omega)),$$

where d, r^-, and r^+ are defined in (3.8). Equivalently, we have

$$\hat{e}_1^{k+1}(L, \omega) = \rho_0 \hat{e}_1^{k-1}(L, \omega), \quad \hat{e}_2^{k+1}(0, \omega) = \rho_0 \hat{e}_2^{k-1}(0, \omega), \tag{5.17}$$

where the convergence factor $\rho_0 = \rho_0(\omega, p, L, \nu, a, b)$ of the new algorithm with Robin transmission conditions is given by

$$\rho_0(\omega, p, L, \nu, a, b) := \left(\frac{\sqrt{a^2 + 4\nu(b + i\omega)} - p}{\sqrt{a^2 + 4\nu(b + i\omega)} + p}\right)^2 e^{-\frac{\sqrt{a^2+4\nu(b+i\omega)}}{\nu}L}. \tag{5.18}$$

For any frequency ω, we can therefore directly compare the performance of the classical Schwarz waveform relaxation algorithm with the new one with Robin transmission conditions: we have $\rho_0 = (\frac{\sqrt{d}-p}{\sqrt{d}+p})^2 \rho_D$, where ρ_D is the classical convergence factor defined in (3.14). This shows that for each ω we have $|\rho_0| < |\rho_D|$ for $p > 0$. Furthermore, for any $\epsilon > 0$ there exists an ω_ϵ such that

$$\int_{|\omega| > \omega_\epsilon} (1 + \omega^2)^{3/4} |\hat{e}_1^0(L, \omega)|^2 d\omega \leq \epsilon$$

because we assume that $e_1^0(L, \cdot)$ is in $H^{\frac{3}{4}}$. Since $|\rho_0| < 1$, we obtain

$$\|\hat{e}_1^{2k}(L, \cdot)\|^2_{H^{\frac{3}{4}}(0,T)} \leq \epsilon + \int_{|\omega| \leq \omega_\epsilon} (1 + \omega^2)^{\frac{3}{4}} |\rho_0(\omega, p, L, \nu, a, b)|^{2k} |\hat{e}_1^0(L, \omega)|^2 d\omega,$$

and taking the supremum of the convergence factor out of the integral, we have

$$\|\hat{e}_1^{2k}(L, \cdot)\|^2_{H^{\frac{3}{4}}(0,T)} \leq \epsilon + \sup_{|\omega| \leq \omega_\epsilon} |\rho_0(\omega, p, L, \nu, a, b)|^{2k} \|\hat{e}_1^0(L, \cdot)\|^2_{H^{\frac{3}{4}}(0,T)}.$$

A similar estimate for the classical algorithm gives

$$\|\hat{e}_1^{2k}(L, \cdot)\|^2_{H^{\frac{3}{4}}(0,T)} \leq \epsilon + \sup_{|\omega| \leq \omega_\epsilon} |\rho_D(\omega, L, \nu, a, b)|^{2k} \|\hat{e}_1^0(L, \cdot)\|^2_{H^{\frac{3}{4}}(0,T)},$$

which shows that improving the convergence factor on a sufficiently large bounded frequency range improves the overall convergence of the algorithm. The choice of a bounded frequency range is further motivated by the fact that computations are performed on a discretized problem, whose grid cannot carry arbitrarily high frequencies. We carefully analyze how to chose the free parameter p for optimal performance of the algorithm in the next subsections.

5.3. Low frequency approximation for the algorithm with overlap. We have seen that the convergence factor of the new algorithm with Robin transmission conditions is given by (5.18), and any choice of the free parameter $p \geq 0$ is admissible to obtain a well posed algorithm. But how should p be chosen, apart from $p \geq 0$? A simple choice is to use a low frequency approximation of the symbols σ_j, $j = 1, 2$, of the optimal transmission operators given in (4.2), based on a Taylor expansion about $\omega = 0$. This is motivated by the fact that with overlap, $L > 0$, the exponential term in the convergence factor (5.18) is exponentially small for ω large, and hence p should be used to make the transmission conditions effective for ω small. Using a Taylor expansion of the square root $\sqrt{a^2 + 4\nu(b + i\omega)}$ in (4.2) about $\omega = 0$, we find

$$\sqrt{a^2 + 4\nu(b + i\omega)} = \sqrt{a^2 + 4\nu b} + \frac{2\nu}{\sqrt{a^2 + 4\nu b}} i\omega + O(\omega^2), \tag{5.19}$$

and hence the low frequency approximation choice for p in the Robin transmission condition is

$$p = p_T := \sqrt{a^2 + 4\nu b}. \tag{5.20}$$

With this choice, the convergence factor vanishes for $\omega = 0$ and also when ω goes to infinity, since $L > 0$. To further analyze the convergence factor, we introduce a change of variables based on the real part of the square root in the convergence factor (5.18),

$$x := \Re(\sqrt{a^2 + 4\nu(b + i\omega)}). \tag{5.21}$$

In this new variable, the convergence factor (5.18) in modulus becomes

$$R_0(x, p, x_0, L) := |\rho_0| = \frac{(x - p)^2 + x^2 - x_0^2}{(x + p)^2 + x^2 - x_0^2} e^{-\frac{Lx}{\nu}}, \tag{5.22}$$

where $x_0^2 := a^2 + 4\nu b$. Note that $R_0 \geq 0$ by definition, which can also be seen from $x^2 \geq a^2 + 4\nu b = x_0^2$ from the change of variables (5.21). Using now the parameter p_T from the Taylor expansion, we find for the Taylor–Robin method (T0 for Taylor of order 0) the convergence factor in modulus to be

$$R_{T0}(x, x_0, L) := R_0(x, p_T, x_0, L) = \frac{x - x_0}{x + x_0} e^{-\frac{Lx}{\nu}} \geq 0, \qquad x \geq x_0. \tag{5.23}$$

THEOREM 5.9 (T0 performance with overlap). *Let $L > 0$ and $x_0 := \sqrt{a^2 + 4\nu b}$. The convergence factor R_{T0} in* (5.23) *of the overlapping Schwarz waveform relaxation algorithm with Robin transmission conditions* (4.1),(4.6) *and $p = p_T$ from the Taylor low frequency approximation* (5.20) *is for $x_0 \leq x < \infty$ uniformly bounded by*

$$R_{T0}(x, x_0, L) \leq \bar{R}_{T0}(x_0, L) := R_{T0}(\bar{x}, x_0, L) = \frac{\bar{x} - x_0}{\bar{x} + x_0} e^{-\frac{L\bar{x}}{\nu}}, \quad \bar{x} = \sqrt{x_0^2 + \frac{2\nu x_0}{L}}. \tag{5.24}$$

For L small, we have $\bar{R}_{T0}(x_0, L) = 1 - 2\sqrt{\frac{2x_0}{\nu}}\sqrt{L} + O(L)$.

Proof. Taking a derivative of R_{T0} with respect to x shows that there is a unique maximum of R_{T0} for $x_0 \leq x < \infty$ at $\bar{x}$, which leads to the bound given in (5.24). An expansion for L small leads then to the asymptotic result of the theorem. ∎

Now in a numerical calculation, two additional issues come into play: First the frequency parameter ω cannot be arbitrarily high, there is a maximum frequency that can be represented on a grid with spacing Δt, and an estimate for this maximum frequency is $\omega_{\max} = \frac{\pi}{\Delta t}$, the signal that oscillates between ± 1 from grid point to grid point. Second, the overlap L is in general not a fixed quantity, one can afford only a few grid cells overlap, and often $L = \Delta x$. The question therefore arises, if for a particular discretization, which might have to satisfy a stability constraint, the bound on the contraction factor in (5.24) is really relevant, or if the highest frequencies represented on the grid of the particular discretization stay below $\bar{\omega}$ where the maximum of R_{T0} is attained, which corresponds to $\bar{x}$ given in (5.24) in the transformed problem. To answer this question, we need to study for which cases the maximum numerical frequency $\omega_{\max}$ stays below $\bar{\omega}$ or, in the transformed problem, under which conditions

$$x_{\max} = \sqrt{\frac{\sqrt{x_0^4 + 16\nu^2\omega_{\max}^2} + x_0^2}{2}} \tag{5.25}$$

stays below $\bar{x}$ given in (5.24). A direct comparison shows that for

$$L > L_0 := \frac{4\nu x_0}{\sqrt{x_0^4 + 16\nu^2\omega_{\max}^2} - x_0^2} \tag{5.26}$$

the maximum numerical frequency $\omega_{\max} > \bar{\omega}$, and hence the bound given in Theorem 5.9 determines the convergence rate of the algorithm. If however $L \leq L_0$, then the maximum on the numerically relevant convergence factor is attained at $\omega_{\max}$. Numerically, we therefore have

$$R_{T0}(x, x_0, L) \leq \tilde{R}_{T0}(x_0, L) := \begin{cases} R_{T0}(\bar{x}, x_0, L) = \dfrac{\bar{x} - x_0}{\bar{x} + x_0} e^{-\frac{L\bar{x}}{\nu}} & \text{if } L > L_0, \\ R_{T0}(x_{\max}, x_0, L) = \dfrac{x_{\max} - x_0}{x_{\max} + x_0} e^{-\frac{Lx_{\max}}{\nu}} & \text{if } L \leq L_0. \end{cases} \tag{5.27}$$

To obtain a concrete asymptotic result for the case where the overlap L is linked to the space discretization Δx, $L = C_1\Delta x$, and the space discretization Δx is linked to the time discretization Δt by a stability or accuracy constraint, $\Delta t = C_2 \Delta x^{\beta}$, $\beta > 0$, we insert these relations into L_0 and expand to find

$$L_0 = \frac{x_0}{\pi}\Delta t + O(\Delta t^2), \tag{5.28}$$

which leads to the following asymptotic results.

Theorem 5.10 (T0 discrete convergence estimate with overlap). *Let $x_0 := \sqrt{a^2 + 4\nu b}$. If $L = C_1\Delta x$ and $\Delta t = C_2 \Delta x^{\beta}$, then the bound $\tilde{R}_{T0}$ in* (5.27) *on the convergence factor estimate of the discretized overlapping Schwarz waveform relaxation algorithm with Robin transmission conditions* (4.1),(4.6) *and $p = p_T$ from the Taylor low frequency approximation* (5.20) *is for Δx small given by*

$$\tilde{R}_{T0} = \begin{cases} 1 - 2\sqrt{\dfrac{2C_1 x_0}{\nu}}\sqrt{\Delta x} + O(\Delta x) & \text{if } \beta > 1 \text{ or } \beta = 1 \text{ and } \frac{C_1}{C_2} > \frac{x_0}{\pi}, \\ 1 - \dfrac{\sqrt{2}(C_2 x_0 + C_1\pi)}{\sqrt{C_2\pi\nu}}\sqrt{\Delta x} + O(\Delta x) & \text{if } \beta = 1 \text{ and } \frac{C_1}{C_2} \leq \frac{x_0}{\pi}, \\ 1 - x_0\sqrt{\dfrac{2C_2}{\pi\nu}}\Delta x^{\frac{\beta}{2}} + o(\Delta x^{\frac{\beta}{2}}) & \text{if } \beta < 1. \end{cases} \tag{5.29}$$

Proof. Comparing $L = C_1\Delta x$ with the expansion of L_0 given in (5.28) and using that $\Delta t = C_2\Delta x^\beta$, we see that for Δx small the first case in (5.27) corresponds to the first case given in (5.29). The asymptotic bound on the convergence factor then follows by simply expanding for $L = C_1\Delta x$ and Δx small. For the second case, one can set $\beta = 1$ and directly expand the second case of (5.27) to find the result given. For the last case, we first notice that $x_{\max}$ satisfies

$$x_{\max} = \sqrt{2\pi\nu}\frac{1}{\sqrt{\Delta t}} + O(\sqrt{\Delta t}) = \sqrt{\frac{2\pi\nu}{C_2}}\Delta x^{-\frac{\beta}{2}} + O(\Delta x^{\frac{\beta}{2}}), \tag{5.30}$$

which together with $L = C_1\Delta x$ gives for the exponential the expansion

$$e^{-\frac{Lx_{\max}}{\nu}} = 1 - C_1\sqrt{\frac{2\pi}{C_2\nu}}\Delta x^{1-\frac{\beta}{2}} + O(\Delta x^{2-\beta}). \tag{5.31}$$

Multiplying this with the expansion for the coefficient in front of the exponential in (5.27), whose expansion is

$$\frac{x_{\max} - x_0}{x_{\max} + x_0} = 1 - x_0\sqrt{\frac{2C_2}{\pi\nu}}\Delta x^{\frac{\beta}{2}} + O(\Delta x^\beta), \tag{5.32}$$

the result follows. ∎

The preceding theorem shows that for explicit discretizations, which have a stability constraint of the type $\Delta t = C_2\Delta x^2$ for this problem and for which the present algorithm would still be of interest for nonmatching time grids, the optimized Schwarz waveform relaxation algorithm with Robin transmission conditions based on a low frequency approximation and an overlap of the order of the spatial discretization Δx will have an asymptotic convergence factor $1 - O(\sqrt{\Delta x})$, as one could expect from the continuous analysis in Theorem 5.9. This is still true for implicit discretizations, as long as Δt is of the same order as Δx. Once Δt becomes much larger than Δx, however, one can expect the algorithm to converge faster asymptotically because of the last relation in (5.29).

5.4. Optimization of the algorithm with overlap. We investigate now if there exists a better choice for p such that the overall convergence factor is smaller than with the parameter from the low frequency approximation. We will use the label O0 for these methods, which stands for optimized of order 0. We place ourselves first again in the continuous context, where $\omega \in \mathbb{R}$, and thus $\omega_{\max} = \infty$. Later, we will also investigate the discretized case where $\omega_{\max} < \infty$. We introduce a change of variables, which will greatly simplify the analysis of the optimal parameter p. Setting $x := \frac{y\nu}{L}$, $p := \frac{\tilde{p}\nu}{L}$, and $x_0 := \frac{y_0\nu}{L}$ in the convergence factor (5.22), we obtain

$$R_0(y, \tilde{p}, y_0) = \frac{(y - \tilde{p})^2 + y^2 - y_0^2}{(y + \tilde{p})^2 + y^2 - y_0^2}e^{-y}, \tag{5.33}$$

which is now an expression independent of the overlap parameter L and the viscosity parameter ν. The best choice for the parameter $\tilde{p}$ is the one that makes R_0 as small as possible uniformly for all $y \geq y_0$ and is hence the solution of the min-max problem

$$\min_{\tilde{p}}\left(\max_{y\geq y_0} R_0(y, \tilde{p}, y_0)\right) = \min_{\tilde{p}\geq 0}\left(\max_{y\geq y_0}\frac{(y - \tilde{p})^2 + y^2 - y_0^2}{(y + \tilde{p})^2 + y^2 - y_0^2}e^{-y}\right), \tag{5.34}$$

where minimizing over nonnegative $\tilde{p}$ is equivalent to minimizing over all $\tilde{p}$, as one can see from (5.33). Note that $\tilde{p}$ nonnegative is also a requirement for the convergence proof of the algorithm in Theorem 5.8. To analyze the min-max problem (5.34), we need two lemmas.

LEMMA 5.11. *The function $y \mapsto R_0(y,\tilde p,y_0)$ defined in* (5.33) *has a unique local maximum at*

$$\bar y(y_0,\tilde p)=\sqrt{\frac{y_0^2+2\tilde p+\sqrt{d(y_0,\tilde p)}}{2}},\quad d(y_0,\tilde p)=\tilde p(-\tilde p^3-4\tilde p^2+(4+2y_0^2)\tilde p+8y_0^2), \tag{5.35}$$

if $0\le\tilde p<\tilde p_1(y_0)$, where $\tilde p_1(y_0)$ is the unique positive root of $d(y_0,\tilde p)$ for $y_0>0$. If $\tilde p\ge\tilde p_1(y_0)$, then $R_0(y,\tilde p,y_0)$ is a monotonically decreasing function of y.

Proof. A partial derivative of $R_0(y,\tilde p,y_0)$ with respect to y gives

$$\frac{\partial R_0}{\partial y}=-\frac{e^{-y}(4y^4-4(2\tilde p+y_0^2)y^2+(\tilde p^2-y_0^2)(y_0^2-4\tilde p-\tilde p^2))}{((\tilde p+y)^2+y^2-y_0^2)^2},$$

and therefore $R_0(y,\tilde p,y_0)$ can have at most two extrema, $\bar y=\sqrt{(y_0^2+2\tilde p+\sqrt{d(y_0,\tilde p)})/2}$ and $\underline{y}=\sqrt{(y_0^2+2\tilde p-\sqrt{d(y_0,\tilde p)})/2}$, with the discriminant $d(y_0,\tilde p)$ given in (5.35). The larger of the two, $\bar y$, must be a maximum, since $R_0\ge 0$ and R_0 goes to 0 as y goes to ∞. Since the discriminant is positive for small positive $\tilde p$ and is negative for large positive $\tilde p$, it must have by continuity at least one real positive root $\tilde p_1(y_0)>0$, $d(y_0,\tilde p_1)=0$. To prove that this root is unique, we use the derivative of $d(y_0,\tilde p)/\tilde p$ with respect to $\tilde p$, which shows that there are two extrema, one at $r_1=-\frac{1}{3}(4-\sqrt{28+6y_0^2})$ and one at $r_2=-\frac{1}{3}(4+\sqrt{28+6y_0^2})$. The larger one, r_1, must be a maximum, since the discriminant goes to $-\infty$ as $\tilde p$ goes to ∞, and thus r_2 is a minimum. Since r_2 is negative, $\tilde p_1$ is the only positive root of the discriminant, since this latter is still positive for arbitrary small $\tilde p$. For $\tilde p\ge\tilde p_1$, R_0 has no extrema in y and hence decreases monotonically to 0 as y goes to infinity. ∎

LEMMA 5.12. *For fixed $y>y_0$ and $\tilde p>0$, we have $\frac{\partial R_0(y,\tilde p,y_0)}{\partial\tilde p}(\tilde p-\tilde p_2(y))\ge 0$, where $\tilde p_2(y):=\sqrt{2y^2-y_0^2}$.*

Proof. A partial derivative of $R_0(y,\tilde p,y_0)$ with respect to $\tilde p$ gives

$$\frac{\partial R_0}{\partial\tilde p}=-\frac{4e^{-y}y(-\tilde p^2+2y^2-y_0^2)}{((\tilde p+y)^2+y^2-y_0^2)^2},$$

which has only one root in $\tilde p$, $\tilde p_2(y)=\sqrt{2y^2-y_0^2}$, which is positive. For $\tilde p<\tilde p_2$, $\frac{\partial R_0}{\partial\tilde p}$ is negative and hence $R_0(y,\tilde p,y_0)$ decreases when $\tilde p$ increases, whereas for $\tilde p>\tilde p_2$, $R_0(y,\tilde p,y_0)$ increases when $\tilde p$ increases. ∎

THEOREM 5.13 (O0 performance with overlap). *Let $L>0$ and $x_0:=\sqrt{a^2+4\nu b}$. The best performance of the optimized overlapping Schwarz waveform relaxation algorithm at the continuous level with Robin transmission conditions* (4.1),(4.6) *is obtained for $p=p^*:=\frac{\tilde p^*\nu}{L}$, where $\tilde p^*$, the solution of the min-max problem* (5.34), *is for $y_0:=\frac{x_0L}{\nu}<y_c$ given by the unique solution $\tilde p^*\ge y_0$ of the nonlinear equation*

$$R_0(y_0,\tilde p^*,y_0)=R_0(\bar y(y_0,\tilde p^*),\tilde p^*,y_0), \tag{5.36}$$

where $R_0(y,\tilde p,y_0)$ is given in (5.33) *and $\bar y(y_0,\tilde p)$ is given in* (5.35). *For $y_0\ge y_c$, $\tilde p^*$ is given by the unique solution of*

$$y_0=\tilde p^*\sqrt{\frac{\tilde p^*}{(4+\tilde p^*)}}. \tag{5.37}$$

The constant y_c is universal, $y_c = 1.618386576\ldots$, and the convergence factor with the optimal p^* is uniformly bounded by

$$R_0(y, \tilde{p}^*, y_0) \leq \bar{R}_{O0}(y_0, \tilde{p}^*) := R_0(\bar{y}(y_0, \tilde{p}^*), \tilde{p}^*, y_0). \tag{5.38}$$

For L small, we have the asymptotic result

$$p^* = \frac{\tilde{p}^* \nu}{L} \approx (2x_0^2 \nu)^{\frac{1}{3}} L^{-\frac{1}{3}}, \quad \bar{R}_{O0} \approx 1 - \left(\frac{2^5 x_0}{\nu}\right)^{\frac{1}{3}} L^{\frac{1}{3}}. \tag{5.39}$$

Proof. By Lemma 5.12, the optimal $\tilde{p}^* \geq y_0$ since for $\tilde{p} < \tilde{p}_2(y_0) = y_0$, increasing $\tilde{p}$ decreases $R_0(y, \tilde{p}, y_0)$ for all $y > y_0$. Now Lemma 5.11 implies that for $y_0 \leq \tilde{p} \leq \tilde{p}_1(y_0)$, the maximum of R_0 in the min-max problem can be attained at $y = y_0$ or at the interior maximum at $\bar{y}$ given in (5.35). For $\tilde{p} = y_0$, we have $R_0(y_0, \tilde{p}, y_0) = R_0(y_0, y_0, y_0) = 0$ and $d(y_0, \tilde{p}) = d(y_0, y_0) = y_0^2(2 + y_0)^2 \geq 0$, and hence $R_0(y, \tilde{p}, y_0)$ has for $y \geq y_0$ a unique maximum at $\bar{y}(y_0, y_0) = \sqrt{y_0(2 + y_0)} > y_0$. Increasing $\tilde{p}$ from y_0 increases $R_0(y_0, \tilde{p}, y_0)$ by Lemma 5.12 monotonically for all $\tilde{p} > \tilde{p}_2(y_0) = y_0$. Increasing $\tilde{p}$ from y_0 also decreases $R_0(\bar{y}(y_0, \tilde{p}), \tilde{p}, y_0)$ by Lemma 5.12, as long as it exists, $\tilde{p} < \tilde{p}_1(y_0)$ according to Lemma 5.11, and $\tilde{p} < \tilde{p}_2(\bar{y}(y_0, \tilde{p})) = \sqrt{2\bar{y}^2 - y_0^2}$, after which $R_0(\bar{y}, \tilde{p}, y_0)$ will increase again according to Lemma 5.12. By continuity, the maximum of R_0 is minimized either for $\tilde{p}_1^*$ satisfying

$$R_0(y_0, \tilde{p}_1^*, y_0) = R_0(\tilde{y}, \tilde{p}_1^*, y_0), \tag{5.40}$$

provided that $\tilde{p}_1^* \leq \tilde{p}_2(\bar{y}(y_0, \tilde{p}_1^*)) = \sqrt{2(\bar{y}(y_0, \tilde{p}_1^*))^2 - y_0^2}$, or for $\tilde{p}_2^*$ given by

$$\tilde{p}_2^* = \tilde{p}_2(\bar{y}(y_0, \tilde{p}_2^*)) = \sqrt{2(\bar{y}(y_0, \tilde{p}_2^*))^2 - y_0^2}. \tag{5.41}$$

It depends on the only parameter left, y_0, which of these two cases is the solution. Imposing $\tilde{p}_1^* = \tilde{p}_2^*$ and both (5.40) and (5.41), we can solve for the value of y_0 where both are equally optimal. We find

$$y_0 = y_c = 1.618386576\ldots, \quad \tilde{p}_1^* = \tilde{p}_2^* = 2.583490822\ldots.$$

Hence for $y_0 < y_c$ and for $y_0 \geq y_c$ (5.40) and (5.41), respectively, give the solution. (5.41) can be simplified by solving it for y_0, which gives (5.37) stated in the theorem, and a derivative with respect to $\tilde{p}^*$ shows that there is a unique positive root $\tilde{p}^*$ for $y_0 > 0$.

The uniform bound given in (5.38) is a direct consequence of (5.36) and (5.37), since in both cases the maximum is attained at $\bar{y}$.

To show the asymptotic result (5.39), we note that for L small and the other problem parameters a, b, and ν fixed, we have y_0 small, since from the variable transform we have $y_0 = \frac{x_0}{\nu} L = \frac{\sqrt{a^2+4\nu b}}{\nu} L$ and therefore the first result (5.36) applies asymptotically, $y_0 < y_c$. To solve (5.36) asymptotically, we insert the ansatz $\tilde{p}^* = C_p y_0^\alpha$ into (5.36) and expand both sides for y_0 small. Using that $\tilde{p}^* \geq y_0$, we find from its definition that asymptotically $\bar{y} \approx \sqrt{2C_p} y_0^{\frac{\alpha}{2}}$. Using this in equation (5.36), we find for the leading order terms

$$1 - 2^{\frac{5}{3}} y_0^{1-\alpha} - y_0 + 2^{\frac{5}{3}} y_0^{1-\alpha+1} y_0 + \cdots = 1 - 2^{\frac{3}{2}} \sqrt{C_p} y_0^{\frac{\alpha}{2}} + 4C_p y_0^\alpha + \cdots,$$

which implies $1 - \alpha = \frac{\alpha}{2}$ and thus $\alpha = \frac{2}{3}$ and $C_p = 2^{\frac{1}{3}}$, which leads to (5.39). ∎

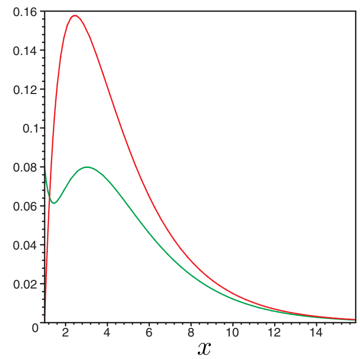

FIG. 5.1. *The top curve is the convergence factor R_{T0} from the Taylor low frequency approximation, and the curve below is the optimized convergence factor R_{O0}, for an example from the numerical section.*

TABLE 5.1
Comparison of the optimal $\tilde{p}^*$ from Theorem 5.13 and its asymptotic approximation.

y_0	0.1	0.01	0.001	0.0001	0.00001
$\tilde{p}^*$	0.2936	0.05952	0.01265	0.002717	0.0005849
Asymptotic $\tilde{p}^*$	0.2714	0.05848	0.01260	0.002714	0.0005848

In Figure 5.1 we show the convergence factors R_{T0} and R_{O0} for an example with $x_0 = 1$, $L = 0.08$, and $\nu = 0.2$ from the numerical section. One can see the better performance of the optimized Robin transmission conditions over the Taylor transmission conditions and also the equioscillation of the optimal choice.

Table 5.1 gives a comparison of the optimal $\tilde{p}^*$ from (5.36) with the asymptotic approximation (5.39). One can see that the asymptotic approximation is very close to the optimal $\tilde{p}^*$ already for moderately small values of y_0, which corresponds to a small overlap L. For larger values of y_0, the asymptotic approximation can be a valuable initial guess for the nonlinear equation solver to find the optimal $\tilde{p}^*$ from (5.36).

In Figure 3.1 on the right, we show the first few iterations, at the end of the time interval, of the optimized Schwarz waveform relaxation algorithm with Robin transmission conditions for the same model problem for which the iterates of the classical Schwarz waveform relaxation algorithm are shown on the left. One can clearly see that the new algorithm with Robin transmission conditions converges much faster than the algorithm with Dirichlet transmission conditions.

Theorem 5.13 gives the optimal choice for the parameter in the Robin transmission conditions of the optimized Schwarz waveform relaxation algorithm at the continuous level. In a numerical setting, however, not all the frequencies are present, as we have seen, and we have to address the question again if the maximum of the convergence factor attained at $\bar{y}$ is relevant in a computation. Letting $L = C_1\Delta x$ and $\Delta t = C_2\Delta x^\beta$ as before, the maximum numerical frequency we can expect on the time discretization grid is $\omega_{\max} = \frac{\pi}{\Delta t} = \frac{\pi}{C_2\Delta x^\beta}$, which corresponds with the variable transform to

(5.42)

$$y_{\max} = \frac{Lx_{\max}}{\nu} = C_1\Delta x\sqrt{\frac{\sqrt{x_0^4 + \left(\frac{4\nu\pi}{C_2\Delta x^\beta}\right)^2} + 2x_0^2}{2}} = C_1\sqrt{\frac{2\pi}{\nu C_2}}\Delta x^{1-\frac{\beta}{2}} + O(\Delta x^{1+\frac{\beta}{2}}),$$

whereas $\bar{y}$ from the optimization in (5.36) has the expansion

$$\bar{y} = 2^{\frac{2}{3}} \left(\frac{x_0 C_1}{\nu}\right)^{\frac{1}{3}} \Delta x^{\frac{1}{3}} + O(\Delta x). \tag{5.43}$$

Hence, if $1 - \frac{\beta}{2} = \frac{1}{3}$ or $\beta = \frac{4}{3}$ and $C_1 = \sqrt{x_0}\left(\frac{2\nu C_2^3}{\pi^3}\right)^{\frac{1}{4}} =: C_c$, the numerical $y_{\max}$ and $\bar{y}$ from the optimization are asymptotically at the same location, which represents the boundary between the usefulness of the continuous optimization result (5.36) and a different optimization for the discretized algorithm, which we show in the following theorem.

THEOREM 5.14 (O0 discrete convergence estimate with overlap). *Let $x_0 := \sqrt{a^2 + 4\nu b}$. If $L = C_1\Delta x$ and $\Delta t = C_2 \Delta x^{\beta}$, then the convergence factor $R_0(y, \tilde{p}, y_0)$ of the discretized overlapping Schwarz waveform relaxation algorithm with Robin transmission conditions* (4.1),(4.6) *is for Δx small bounded for all $y \in [y_0, y_{\max}]$ by $\tilde{R}_{O0}$, where $y_{\max}$ is given in* (5.42)*, $\tilde{R}_{O0}$ and $\tilde{p}^*$ satisfy*

(5.44)

$$\begin{aligned}
&\tilde{R}_{O0} \approx 1 - \left(\tfrac{2^5 C_1 x_0}{\nu}\right)^{\frac{1}{3}} \Delta x^{\frac{1}{3}},\ p^* \approx \left(\tfrac{2x_0^2\nu}{C_1}\right)^{\frac{1}{3}} \Delta x^{-\frac{1}{3}} && \text{if } \beta > \tfrac{4}{3} \text{ or } \beta = \tfrac{4}{3} \text{ and } C_1 > C_c,\\
&\tilde{R}_{O0} \approx 1 - \tfrac{8C_1x_0}{C_p\nu}\Delta x^{\frac{1}{3}}, \qquad p^* \approx C_p \Delta x^{-\frac{1}{3}} && \text{if } \beta = \tfrac{4}{3} \text{ and } C_1 \le C_c,\\
&\tilde{R}_{O0} \approx 1 - 2\left(\tfrac{2C_2x_0^2}{\nu\pi}\right)^{\frac{1}{4}} \Delta x^{\frac{\beta}{4}},\ p^* \approx \left(\tfrac{2^3 x_0^2 \nu\pi}{C_2}\right)^{\frac{1}{4}} \Delta x^{-\frac{\beta}{4}} && \text{if } \beta < \tfrac{4}{3},
\end{aligned}$$

and the constants are given by $C_p = \frac{1}{2C_2}(\sqrt{\pi^2 C_1^2 + 8\sqrt{2\nu\pi} C_2^{\frac{3}{2}} x_0} - \pi C_1)$, $C_c = \sqrt{x_0}(\frac{2\nu C_2^3}{\pi^3})^{\frac{1}{4}}$.

Proof. The first case is a direct consequence of Theorem 5.13, which applies in this case, since the maximum $\bar{y}$ is relevant for the numerical discretization if $\beta > \frac{4}{3}$ or $\beta = \frac{4}{3}$ and $C_1 > C_c$, as can be see from (5.42) and (5.43). For case two and three, the local maximum at $\bar{y}$ lies outside of the numerical frequencies, $\bar{y} > y_{\max}$, and hence the min-max problem needs to be adapted to this situation; the maximum needs now be minimized only for $y \in [y_0, y_{\max}]$. For a small overlap, which corresponds to y_0 small, the solution is achieved according to Theorem 5.13 when

$$R_0(y_0, \tilde{p}^*, y_0) = R_0(y_{\max}, \tilde{p}^*, y_0). \tag{5.45}$$

Expanding both sides asymptotically for small Δx, we find, using the ansatz $\tilde{p}^* = \tilde{C}_p \Delta x^{\alpha}$, that the leading order terms of (5.45) are

$$1 - \frac{4x_0C_1}{\tilde{C}_p\nu}\Delta x^{1-\alpha} + \cdots = 1 - \frac{2\tilde{C}_p}{C_1}\sqrt{\frac{\nu C_2}{2\pi}}\Delta x^{\alpha - 1 + \frac{\beta}{2}} - C_1\sqrt{\frac{2\pi}{\nu C_2}}\Delta x^{1-\frac{\beta}{2}} + \cdots .$$

Hence in the limiting case, where $\beta = \frac{4}{3}$, we have $\alpha = \frac{2}{3}$, and both terms on the right have the same exponent. This leads to the constant C_p given in the theorem in case two, after having used the back transform $p^* = \frac{\tilde{p}^*\nu}{C_1\Delta x}$. If however $\beta < \frac{4}{3}$, then the last term on the right-hand side is of lower order. Balancing the remaining two, we find for the exponents $1 - \alpha = \alpha - 1 + \frac{\beta}{2}$ or $\alpha = 1 - \frac{\beta}{4}$ and the constant $\tilde{C}_p = C_1\left(\frac{2^3 x_0^2 \pi}{\nu^3 C_2}\right)^{\frac{1}{4}}$, which leads after the back transform to the last case stated in the theorem. ∎

5.5. Convergence for the nonoverlapping algorithm. We now assume that the overlap is zero, $L = 0$. We first analyze convergence of the algorithm in the appropriate Sobolev spaces. The convergence analysis for the nonoverlapping case is based on energy estimates and follows an idea from [28], which has also been used in [6] and [33] for Schwarz algorithms applied to steady problems and in [11] for a nonoverlapping Schwarz waveform relaxation algorithm for hyperbolic evolution equations.

THEOREM 5.15. *Without overlap, $L = 0$, the Schwarz waveform relaxation algorithm* (4.1),(4.6) *converges for $p > 0$ in $(L^\infty(0,T;L^2(\Omega_1))) \cap L^2(0,T;H^1(\Omega_1)) \times (L^\infty(0,T;L^2(\Omega_2))) \cap L^2(0,T;H^1(\Omega_2))$ to the solution u of* (2.1) *for any initial guess $g_0 \in H^{\frac{1}{4}}(0,T)$ and $g_L \in H^{\frac{1}{4}}(0,T)$.*

Proof. As in the proof of Theorem 5.2 we obtain the energy estimates

$$\frac{1}{2}\frac{d}{dt}\|e_1^k\|^2 + \nu\|\partial_x e_1^k\|^2 + b\|e_1^k\|^2 - \left(\nu\partial_x e_1^k - \frac{a}{2}e_1^k\right)(0)e_1^k(0) = 0, \tag{5.46}$$

$$\frac{1}{2}\frac{d}{dt}\|e_2^k\|^2 + \nu\|\partial_x e_2^k\|^2 + b\|e_2^k\|^2 + \left(\nu\partial_x e_2^k - \frac{a}{2}e_2^k\right)(0)e_2^k(0) = 0. \tag{5.47}$$

Introducing the boundary operators $\mathcal{B}^+ = \partial_x + \mathcal{S}_1$, $\mathcal{B}^- = \partial_x + \mathcal{S}_2$ and rewriting the terms on the interface in the form $(\nu\partial_x e - \frac{a}{2}e)e = \frac{\nu^2}{2p}((\mathcal{B}^+e)^2 - (\mathcal{B}^-e)^2)$, we obtain the new energy estimates

$$\frac{1}{2}\frac{d}{dt}\|e_1^k\|^2 + \nu\|\partial_x e_1^k\|^2 + b\|e_1^k\|^2 + \frac{\nu^2}{2p}(\mathcal{B}^-e_1^k)^2(0) = \frac{\nu^2}{2p}(\mathcal{B}^+e_1^k)^2(0), \tag{5.48}$$

$$\frac{1}{2}\frac{d}{dt}\|e_2^k\|^2 + \nu\|\partial_x e_2^k\|^2 + b\|e_2^k\|^2 + \frac{\nu^2}{2p}(\mathcal{B}^+e_2^k)^2(0) = \frac{\nu^2}{2p}(\mathcal{B}^-e_2^k)^2(0). \tag{5.49}$$

Now note that the transmission conditions can be expressed with the operators $\mathcal{B}^\pm$,

$$\mathcal{B}^+e_1^k = \mathcal{B}^+e_2^{k-1}, \quad \mathcal{B}^-e_2^k = \mathcal{B}^-e_1^{k-1} \text{ on } \{0\} \times (0,T).$$

Replacing the corresponding terms in the two equations (5.48) and (5.49), adding the resulting equations, and summing in k, we get a telescopic sum on the interfaces and therefore

$$\begin{aligned} &\textstyle\sum_{k=1}^K \left[\frac{1}{2}\frac{d}{dt}(\|e_1^k\|^2 + \|e_2^k\|^2) + \nu(\|\partial_x e_1^k\|^2 + \|\partial_x e_2^k\|^2) + b(\|e_1^k\|^2 + \|e_2^k\|^2)\right] + \\ &\qquad + \tfrac{\nu^2}{2p}((\mathcal{B}^-e_1^K)^2 + (\mathcal{B}^+e_2^K)^2)(0) = \tfrac{\nu^2}{2p}((\mathcal{B}^-e_1^1)^2 + (\mathcal{B}^+e_2^1)^2)(0). \end{aligned} \tag{5.50}$$

We can now integrate in time, and since the initial values of the error vanish, the sum of the energies over all the iterates remains bounded. Hence the energy in the iterates needs to go to zero and the algorithm converges. ∎

5.6. Low frequency approximation for the algorithm without overlap. One can choose the free parameter p in the Robin transmission conditions based on a low frequency Taylor approximation, as given in (5.20). But now there is no overlap to be effective on the high frequencies, the convergence factor (5.23) becomes

$$R_{T0}(x, x_0) = \frac{x - x_0}{x + x_0}, \tag{5.51}$$

where $x \geq x_0$ is given by the variable transform (5.21). Clearly R_{T0} is a monotonically increasing function of x for $x \geq x_0$ and tends to one as x tends to infinity. There is therefore no uniform bound on R_{T0} which is strictly less than one for all $x \geq x_0$ in the case without overlap. But we have already seen that in a numerical calculation the frequency parameter ω cannot be arbitrarily high. It suffices therefore for the numerical case to find a bound for R_{T0} for $x_0 \leq x \leq x_{\max}$, where $x_{\max}$ is given in (5.25).

THEOREM 5.16 (T0 discrete convergence estimate without overlap). *Let $x_0 := \sqrt{a^2 + 4\nu b}$ and $L = 0$. Then the convergence factor estimate R_{T0} of the discretized nonoverlapping Schwarz waveform relaxation algorithm with Robin transmission conditions* (4.1),(4.6) *and $p = p_T$ from the Taylor low frequency approximation* (5.20) *is for $x_0 \leq x \leq x_{\max}$, where $x_{\max}$ is defined in* (5.25), *bounded by*

(5.52)
$$R_{T0}(x, x_0) \leq \tilde{R}_{T0}(x_0) := R_{T0}(x_{\max}, x_0) = \frac{\sqrt{\sqrt{\Delta t^2 x_0^4 + 16\nu^2\pi^2} + x_0^2\Delta t} - \sqrt{2\Delta t}x_0}{\sqrt{\sqrt{\Delta t^2 x_0^4 + 16\nu^2\pi^2} + x_0^2\Delta t} + \sqrt{2\Delta t}x_0}.$$

For Δt small, we have $\tilde{R}_{T0}(x_0) = 1 - x_0\sqrt{\frac{2}{\nu\pi}}\sqrt{\Delta t} + O(\Delta t)$.

Proof. By the monotonicity of R_{T0} in x, the bound for $x_0 \leq x \leq x_{\max}$ on R_{T0} is attained at $x = x_{\max}$, which leads, using the variable transform (5.21) and $\omega_{\max} = \frac{\pi}{\Delta t}$, to the bound given in (5.52). □

Now we can compare the asymptotic performance of the algorithm without overlap to the performance of the algorithm with overlap. If in the discretization the time step Δt is linked to the spatial discretization step Δx by the relation $\Delta t = C_2\Delta x^\beta$, then we see by comparing the results of Theorem 5.10 with the results of Theorem 5.16 that for $\beta \geq 1$ adding an overlap of size Δx does improve the asymptotic performance of the algorithm, whereas for $\beta < 1$ adding an overlap of the order of Δx does not improve the asymptotic performance. In particular this shows that with the Taylor transmission conditions and using an explicit time discretization with the stability constraint $\Delta t = C_1\Delta x^2$, an overlap is helpful. Note that if an explicit scheme is used with the same time steps in both subdomains, there is no need to iterate, since one can explicitly advance the algorithm on the interface as well. A subdomain iteration would still be of interest if one uses nonmatching time grids, however, see, for example, [11].

5.7. Optimization of the algorithm without overlap. As in the case with overlap, there is a better choice for p than the low frequency approximation based on a Taylor expansion of the optimal symbol. We can again try to choose p such that the convergence factor

$$R_0(x, p, x_0) = \frac{(x-p)^2 + x^2 - x_0^2}{(x+p)^2 + x^2 - x_0^2} \tag{5.53}$$

is minimized over all $x_0 \leq x \leq x_{\max}$. Hence the optimal choice for p for the discretized algorithm is the solution of the min-max problem

$$\min_p \left(\max_{x_0 \leq x \leq x_{\max}} R_0(x, p, x_0) \right) = \min_{p \geq 0} \left(\max_{x_0 \leq x \leq x_{\max}} \frac{(x-p)^2 + x^2 - x_0^2}{(x+p)^2 + x^2 - x_0^2} \right), \tag{5.54}$$

where minimizing over nonnegative p is equivalent to minimizing over all p, as one can see from (5.53). The following theorem can be proved as in the case with overlap.

THEOREM 5.17 (O0 performance without overlap). *Let $L = 0$, $x_0 := \sqrt{a^2 + 4\nu b}$, and $x_{\max} < \infty$ be given. Then the best performance of the optimized nonoverlapping*

TABLE 5.2

Summary of the asymptotic convergence factors for the various parameter choices in the Robin transmission conditions for $\Delta t = \Delta x^\beta$.

Method	Convergence factor	Parameter p
T0 overlap Δx	$\begin{cases} 1-O(\sqrt{\Delta x}) & \text{if } \beta \geq 1 \\ 1-O(\Delta x^{\frac{\beta}{2}}) & \text{if } \beta < 1 \end{cases}$	$\sqrt{a^2+4\nu b}$
O0 overlap Δx	$\begin{cases} 1-O(\Delta x^{\frac{1}{3}}) & \text{if } \beta \geq \frac{4}{3} \\ 1-O(\Delta x^{\frac{\beta}{4}}) & \text{if } \beta < \frac{4}{3} \end{cases}$	$(2\nu(a^2+4\nu b))^{\frac{1}{3}}\Delta x^{-\frac{1}{3}}$ $(8\nu\pi(a^2+4\nu b))^{\frac{1}{4}}\Delta x^{-\frac{\beta}{4}}$
T0 no overlap	$1-O(\sqrt{\Delta t})$	$\sqrt{a^2+4\nu b}$
O0 no overlap	$1-O(\Delta t^{\frac{1}{4}})$	$(8\nu\pi(a^2+4\nu b))^{\frac{1}{4}}\Delta t^{-\frac{1}{4}}$

Schwarz waveform relaxation algorithm with Robin transmission conditions (4.1),(4.6) *is obtained for $p = p^*$, where p^*, the solution of the min-max problem* (5.54), *is for $x_{\max} \geq \frac{1+\sqrt{5}}{2}x_0$ given by*

(5.55)

$$p^* = \sqrt{x_0(2x_{\max}+x_0)}, \qquad R_0(x,p^*,x_0) \leq \tilde{R}_{O0} = \frac{x_{\max}+x_0-\sqrt{2x_{\max}x_0+x_0^2}}{x_{\max}+x_0+\sqrt{2x_{\max}x_0+x_0^2}},$$

and for $x_{\max} < \frac{1+\sqrt{5}}{2}x_0$ we have

$$p^* = \sqrt{2x_{\max}^2-x_0^2}, \qquad R_0(x,p^*,x_0) \leq \tilde{R}_{O0} = \frac{\sqrt{2x_{\max}^2-x_0^2}-x_{\max}}{\sqrt{2x_{\max}^2-x_0^2}+x_{\max}}. \tag{5.56}$$

THEOREM 5.18 (O0 discrete convergence estimate without overlap). *Let $L = 0$ and $x_0 := \sqrt{a^2+4\nu b}$. If the nonoverlapping Schwarz waveform relaxation algorithm with optimized Robin transmission conditions is discretized in time with time step Δt, then for Δt small we have*

$$p^* = (2^3x_0^2\pi\nu)^{\frac{1}{4}}\Delta t^{-\frac{1}{4}} + O(\Delta t^{\frac{1}{4}}), \quad \tilde{R}_{O0} = 1-2\left(\frac{2x_0^2}{\pi\nu}\right)^{\frac{1}{4}}\Delta t^{\frac{1}{4}} + O(\sqrt{\Delta t}). \tag{5.57}$$

Proof. Using the variable transform (5.21), $x_{\max}$ behaves like

$$x_{\max} = \sqrt{2\pi\nu}\Delta t^{-\frac{1}{2}} + O(\sqrt{\Delta t}),$$

and thus the first result of (5.55) in Theorem 5.17 applies. Expanding p^* and $\tilde{R}_{O0}$ from (5.55) leads to (5.57). ∎

To summarize the results of this section, we show in Table 5.2 an overview of the performance one can obtain with the various choices of the parameter p in the transmission conditions. It is interesting to note that, for optimized Schwarz waveform relaxation methods, without overlap does not necessarily mean less performance than with overlap: in the T0 case, if $\beta \leq 1$, the performance of the overlapping and nonoverapping algorithms is the same, and the same holds in the O0 case if $\beta \leq \frac{4}{3}$.

6. Numerical results. We perform in this section numerical experiments to measure the convergence factors of the numerical implementation of the Schwarz waveform relaxation algorithms analyzed at the continuous level in this paper. We use the parabolic model problem (2.1) with $\Omega = (0,6)$. We impose homogeneous boundary conditions, $u(0,t) = 0$ and $u(6,t) = 0$, and use various initial conditions

$u(x, 0)$, $x \in \Omega$. We first use a decomposition of the domain Ω into the two subdomains $\Omega_1 = (0, L_2)$ and $\Omega_2 = (L_1, 6)$, $L_1 \leq L_2$, and hence $L = L_2 - L_1$. We refer with the term iteration to a double iteration of the respective algorithms, since for two subdomains one can perform all the iterations in an alternating fashion and thus obtain the even iterates on one subdomain and the odd ones on the other without having to compute the remaining ones. We show results of numerical experiments for only the algorithm with overlap since with overlap we can compare the results to the classical Schwarz waveform relaxation algorithm with Dirichlet transmission conditions, which does not converge without overlap.

6.1. Dirichlet transmission conditions. In this first set of experiments, we use the classical Schwarz waveform relaxation algorithm with Dirichlet transmission conditions analyzed in section 3. We chose for the problem parameters $\nu = 0.2$, $a = 1$, and $b = 0$. We discretize (2.1) using an upwind finite difference discretization in space with mesh parameter $\Delta x = 0.02$ and a backward Euler discretization in time, with time step $\Delta t = 0.005$. We chose $L_1 = 2.96$ and $L_2 = 3.04$, which means the overlap is $L = 0.08$, and we compute the numerical solution in the time interval $[0, T]$. Using as initial condition

$$u(x, 0) = e^{-3(1.2-x)^2},$$

we have already shown in Figure 3.1 for this example the first few iterations at the end of the time interval $T = 2.5$, where we started the algorithm with a zero initial guess. We show in Figure 6.1 the convergence behavior of the classical Schwarz waveform relaxation algorithm for this example for three different lengths of the time interval, $T = 1$, $T = 2.5$, and $T = 10$, together with the linear bound on the convergence rate from Theorem 3.3 and the superlinear convergence bound from Theorem 3.4. The dashed curve shows the error measured in the L_2 norm between the converged solution and the iterates at the interface L_2. One can clearly see that the behavior of the algorithm depends on the length of the time interval T, as predicted by Theorem 3.4. For short time intervals, the superlinear bound on the convergence rate is sharper, and hence the algorithm must converge superlinearly, as shown in Figure 6.1 on the left. If the time interval becomes longer, as in the middle graph of Figure 6.1, the linear bound of Theorem 3.3 is sharper than the superlinear one early in the iteration, and hence the algorithm converges linearly. But later the superlinear bound becomes sharper and hence a transition to the superlinear convergence regime occurs. For long time intervals, the initial linear convergence regime also prevails for more iterations, as one can see in Figure 6.1 on the right.

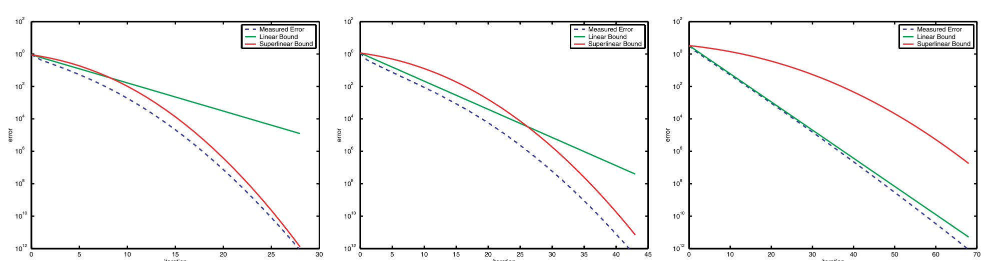

FIG. 6.1. *Convergence rate of the classical Schwarz waveform relaxation algorithm with Dirichlet transmission conditions together with the theoretical linear and superlinear bounds on the convergence rates: on the left for $T = 1$, in the middle for $T = 2.5$, and on the right for $T = 10$.*

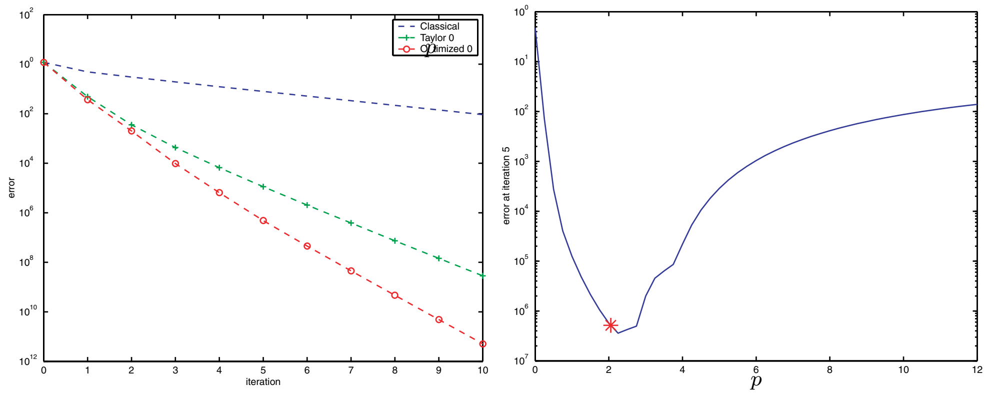

Fig. 6.2. *Left: Convergence rates of the classical Schwarz waveform relaxation algorithm with Dirichlet transmission conditions compared to the same algorithm with the new Robin transmission conditions, with the low frequency Taylor approximation or optimized. Right: The error obtained running the algorithm with Robin transmission conditions for* 5 *steps and various choices of the free parameter p, and indicated by a star the choice p^* predicted by the theory.*

6.2. Robin transmission conditions. We now change the transmission conditions in the Schwarz waveform relaxation algorithm to Robin transmission conditions. Using the same numerical configuration as in the previous subsection, we obtain for the parameter p in the transmission conditions using a Taylor expansion $p = p_T = 1$, and using the optimization from Theorem 5.13, we obtain $p = p^* = 2.054275607$. In Figure 3.1 on the right, we have already seen the first few iterations at the end of the time interval $T = 2.5$ for this example with the optimal parameter p^*, starting the iteration with the zero initial guess. In Figure 6.2 one can see how much faster the algorithm converges with Robin transmission conditions compared to the classical algorithm. One can also see that the optimized parameter p^* leads to an even better performance than the parameter p_T from the Taylor transmission conditions. Note that for all the results comparing the performance of the algorithms, we started the iteration with a random initial guess. This is important to obtain a relevant comparison since, for smooth solutions starting with a smooth initial guess, high frequencies would not be present on the mesh and thus a much coarser mesh would have been sufficient for the computation. The random initial guess has the effect that the mesh resolution is indeed needed to resolve the iteration and thus corresponds to the relevant case in practice.

Next, we verify if the optimal choice for the parameter $p = p^*$ derived using the continuous Fourier analysis in Theorem 5.13 really corresponds to the best choice one can make in the fully discretized algorithm. In Figure 6.2 on the right we show the error obtained after running the Schwarz waveform relaxation algorithm with Robin transmission conditions for five steps using various values for the free parameter p in the transmission conditions. The optimal choice p^* from Theorem 5.13 is indicated by a star. Clearly the continuous analysis predicts the optimal choice of the parameter p very well.

Finally, we illustrate the asymptotic analysis by performing two sets of experiments according to Theorems 5.10 and 5.14. We choose the same problem parameters as before but start now with a coarser mesh both in space and time, $\Delta x = 0.08$ and $\Delta t = 0.02$, and we fix the overlap to be $L = \Delta x$. We then run the classical and optimized Schwarz waveform relaxation algorithms with Taylor–Robin transmission

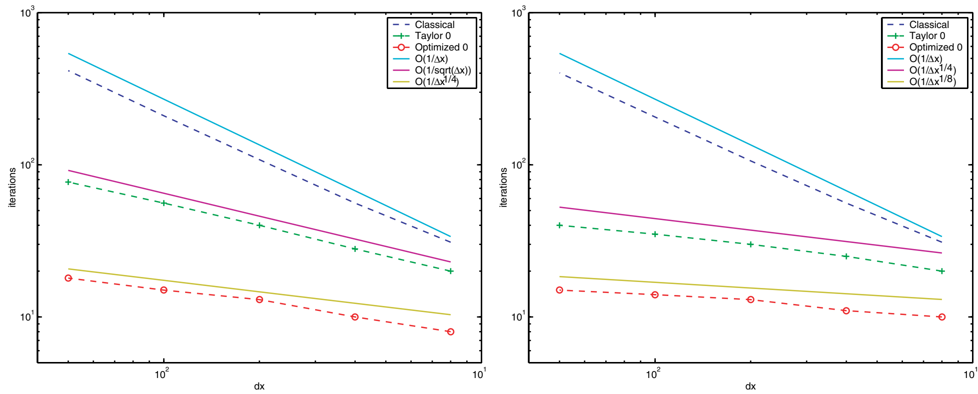

FIG. 6.3. *Asymptotic behavior as the mesh is refined with an overlap $L = \Delta x$: on the left the case where $\Delta t = O(\Delta x)$ and on the right the case where $\Delta t = O(\sqrt{\Delta x})$, together with the predicted rates from the analysis, both for the classical and the optimized Schwarz waveform relaxation algorithms with Taylor and optimized Robin transmission conditions.*

conditions until the error becomes smaller than 10^{-6} and count the number of iterations. We repeat this experiment dividing Δx and Δt by 2 several times, which implies $\Delta t = O(\Delta x)$. This corresponds to (3.16) for the classical algorithm, where the convergence factor should behave like $1 - O(\Delta x)$. For the algorithm with Taylor transmission conditions it corresponds to the case in Theorem 5.10, where the convergence factor should behave like $1 - O(\sqrt{\Delta x})$, and for the optimized algorithm it corresponds to the case in Theorem 5.14, where the convergence factor should behave like $1 - O(\Delta x^{\frac{1}{4}})$. Figure 6.3 shows on the left the results obtained from these experiments. One can see that the asymptotic analysis predicts very well the numerical behavior of the algorithms. Next, we perform a similar experiment, starting with the same values for Δx and Δt, but now we divide Δx by 2 each time and Δt only by $\sqrt{2}$ (such a refinement is admissible since our scheme is implicit), which implies $\Delta t = O(\sqrt{\Delta x})$. While this does not change anything for the classical algorithm, which still has the same bad convergence factor $1 - O(\Delta x)$, for the algorithm with Taylor–Robin transmission conditions now case 3 of Theorem 5.10 applies, and the algorithm should show the much better convergence factor $1 - O(\Delta x^{\frac{1}{4}})$. The optimized algorithm has according to Theorem 5.14 now the even better convergence factor $1 - O(\Delta x^{\frac{1}{8}})$, virtually independent of Δx. In Figure 6.3 on the right, one can clearly see that this is the case. The algorithm has different asymptotic convergence factors with the same overlap, depending on the discretization in time, as predicted.

6.3. Experiments with many subdomains. We now show experiments which indicate that the results we obtained for two subdomains are also relevant for many subdomains. Using the same model problem as before, we now decompose the domain into eight subdomains. In Figure 6.4, we show in the top row the first three iterations of the classical Schwarz waveform relaxation algorithm, and below we show the same iterations for the algorithm with optimized Robin transmission conditions. This clearly shows how important the transmission conditions are in the many subdomain case. We show the corresponding convergence rates in Figure 6.5 on the left, and on the right we perform the same asymptotic experiments as in Figure 6.3 on the left but now with eight subdomains, which indicates that the results of Theorems 5.10 and 5.14 also hold for more than two subdomains.

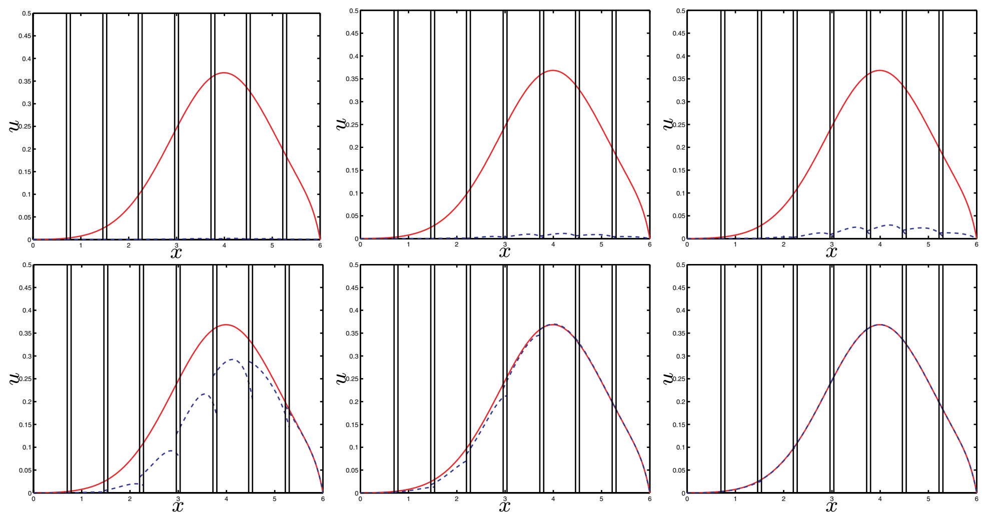

FIG. 6.4. *From left to right, the first iterates $u_j^k(x,T)$, $j = 1, \ldots, 8$, (dashed) at the end of the time interval $t = T$ together with the exact solution (solid) for the same model problem as before: top row the classical algorithm and bottom row the optimized algorithm.*

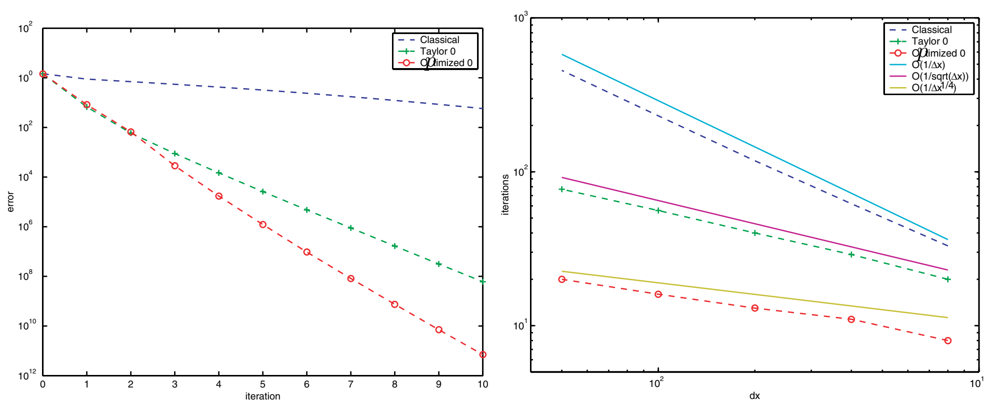

FIG. 6.5. *Left: convergence rate comparison for the eight subdomain case. Right: Asymptotic behavior as the mesh is refined with an overlap $L = \Delta x$ for the eight subdomain case, with $\Delta t = O(\Delta x)$, together with the predicted rates from the two subdomain analysis.*

7. Conclusions. We have analyzed Schwarz waveform relaxation algorithms for advection reaction diffusion equations. We have shown that these methods, using the classical Dirichlet transmission conditions, are well defined and have a convergence rate which is bounded both by a linear and a superlinear rate. Both rates can be sharp, depending on the length of the time interval of the simulation. We then showed that there exist much better transmission conditions than the classical Dirichlet conditions. Optimal transmission conditions are transparent conditions, but they are in general nonlocal and thus less convenient to use. We introduced instead Robin transmission conditions in the Schwarz waveform relaxation algorithm, showed that the new algorithm is well posed and convergent, even if there is no overlap, and analyzed how to chose the free parameter in the new transmission conditions. We also gave asymptotic results when the overlap or the mesh parameters become small.

We finally illustrated our findings with numerical experiments which document the relevance of our continuous analysis.

REFERENCES

[1] K. Burrage, *Parallel and Sequential Methods for Ordinary Differential Equations*, Oxford University Press, New York, 1995.

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

[3] X.-C. Cai, *Multiplicative Schwarz methods for parabolic problems*, SIAM J. Sci. Comput., 15 (1994), pp. 587–603.

[4] P. D'Anfray, L. Halpern, and J. Ryan, *New trends in coupled simulations featuring domain decomposition and metacomputing*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 953–970.

[5] D. S. Daoud and M. J. Gander, *Overlapping Schwarz waveform relaxation for convection reaction diffusion problems*, in Proceedings of the 13th International Conference on Domain Decomposition, 2001, pp. 253–260.

[6] B. Després, *Méthodes de Décomposition de Domaine pour les Problèmes de Propagation d'Ondes en Régimes Harmoniques*, Ph.D. thesis, Paris IX, 1991.

[7] B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp., 31 (1977), pp. 629–651.

[8] B. Engquist and H.-K. Zhao, *Absorbing boundary conditions for domain decomposition*, Appl. Numer. Math., 27 (1998), pp. 341–365.

[9] M. J. Gander and G. H. Golub, *A non-overlapping optimized Schwarz method which converges with an arbitrarily weak dependence on h*, in Proceedings of the Fourteenth International Conference on Domain Decomposition Methods, 2002.

[10] M. J. Gander, L. Halpern, and F. Nataf, *Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation*, in Proceedings of the Eleventh International Conference of Domain Decomposition Methods, 1999.

[11] M. J. Gander, L. Halpern, and F. Nataf, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM J. Numer. Anal., 41 (2003), pp. 1643–1681.

[12] M. J. Gander and L. Halpern, *Méthodes de relaxation d'ondes pour l'équation de la chaleur en dimension* 1, C. R. Acad. Sci. Paris Ser. I, 336 (2003), pp. 519–524.

[13] 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.

[14] M. J. Gander and H. Zhao, *Overlapping Schwarz waveform relaxation for parabolic problems in higher dimension*, in Proceedings of Algoritmy 14, Slovak Technical University, 1997, pp. 42–51.

[15] M. J. Gander, *Analysis of Parallel Algorithms for Time Dependent Partial Differential Equations*, Ph.D. thesis, Stanford University, Stanford, CA, 1997.

[16] M. J. Gander, *Overlapping Schwarz for parabolic problems*, in Proceedings of the Ninth International Conference on Domain Decomposition Methods, 1997, pp. 97–104.

[17] M. J. Gander, *Optimized Schwarz Methods*, Technical report 2003-01, McGill University, Montreal, QC, Canada, 2003.

[18] E. Giladi and H. B. Keller, *Space time domain decomposition for parabolic problems*, Numer. Math., 93 (2002), pp. 279–313.

[19] L. Halpern, *Artificial boundary conditions for the advection-diffusion equations*, Math. Comp., 174 (1986), pp. 425–438.

[20] M. Hu, K. Jackson, J. Janssen, and S. Vandewalle, *Remarks on the optimal convolution kernel for CSOR waveform relaxation*, Adv. Comput. Math., 7 (1997), pp. 135–156.

[21] J. Janssen and S. Vandewalle, *Multigrid waveform relaxation on spatial finite element meshes: The continuous-time case*, SIAM J. Numer. Anal., 33 (1996), pp. 456–474.

[22] J. Janssen and S. Vandewalle, *Multigrid waveform relaxation on spatial finite element meshes: The discrete-time case*, SIAM J. Sci. Comput., 17 (1996), pp. 133–155.

[23] J. Janssen and S. Vandewalle, *On SOR waveform relaxation methods*, SIAM J. Numer. Anal., 34 (1997), pp. 2456–2481.

[24] C. Japhet, *Optimized Krylov-Ventcell method. Application to convection-diffusion problems*, in Proceedings of the 9th International Conference on Domain Decomposition Methods, 1998, pp. 382–389.

[25] R. Jeltsch and B. Pohl, *Waveform relaxation with overlapping splittings*, SIAM J. Sci. Comput., 16 (1995), pp. 40–49.

[26] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. CAD IC Syst., 1 (1982), pp. 131–145.

[27] J.-L. Lions and E. Magenes, *Problèmes aux Limites Non Homogènes et Applications*, Trav. Recherches Math. 17–18, Dunod, Paris, 1968.

[28] P.-L. Lions, *On the Schwarz alternating method*. III: *A variant for nonoverlapping subdomains*, in Proceedings of the Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, Houston, TX, SIAM, Philadelphia, PA, 1990.

[29] C. Lubich and A. Ostermann, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.

[30] C. Lubich and A. Schädle, *Fast convolution for non-reflecting boundary conditions*, SIAM J. Sci. Comput., 24 (2002), pp. 161–182.

[31] V. Martin, *An optimized Schwarz waveform relaxation method for unsteady convection diffusion equation*, Appl. Numer. Math., 52 (2005), pp. 401–428.

[32] G. A. Meurant, *Numerical experiments with a domain decomposition method for parabolic problems on parallel computers*, in Proceedings of the Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, PA, 1991.

[33] F. Nataf, F. Rogier, and E. de Sturler, *Optimal Interface Conditions for Domain Decomposition Methods*, Technical report 301, CMAP (Ecole Polytechnique), 1994.

[34] F. Nataf and F. Rogier, *Factorization of the convection-diffusion operator and the Schwarz algorithm*, M^3AS, 5 (1995), pp. 67–93.

[35] O. Nevanlinna, *Remarks on Picard-Lindelöf iterations part i*, BIT, 29 (1989), pp. 328–346.

[36] R. Rannacher and G. Zhou, *Analysis of a domain-splitting method for nonstationary convection-diffusion problems*, East-West J. Numer. Math., 2 (1994), pp. 151–174.

[37] W. Rudin, *Real and Complex Analysis*, McGraw–Hill, New York, 1966.

[38] S. Vandewalle and G. Horton, *Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods*, Computing, 54 (1995), pp. 317–330.