Numer Algor (2013) 64:221–243
DOI 10.1007/s11075-012-9662-y

ORIGINAL PAPER

Optimization of Schwarz waveform relaxation over short time windows

Yves Courvoisier · Martin J. Gander

Received: 22 December 2011 / Accepted: 29 October 2012 / Published online: 8 December 2012

Abstract Schwarz waveform relaxation algorithms (SWR) are naturally parallel solvers for evolution partial differential equations. They are based on a decomposition of the spatial domain into subdomains, and a partition of the time interval of interest into time windows. On each time window, an iteration, during which subproblems are solved in space-time subdomains, is then used to obtain better and better approximations of the overall solution. The information exchange between subdomains in space-time is performed through classical or optimized transmission conditions (TCs). We analyze in this paper the optimization problem when the time windows are short. We use as our model problem the optimized SWR algorithm with Robin TCs applied to the heat equation. After a general convergence analysis using energy estimates, we prove that in one spatial dimension, the optimized Robin parameter scales like the inverse of the length of the time window, which is fundamentally different from the known scaling on general bounded time windows, which is like the inverse of the square root of the time window length. We illustrate our analysis with a numerical experiment.

Keywords Schwarz waveform relaxation · Optimized Robin transmission conditions · Short time windows · Heat equation

1 Introduction

Waveform relaxation algorithms (WR) were invented in the circuit community, see [21], in order to simulate very large scale circuits on parallel computers. They solve

Y. Courvoisier (✉) · M. J. Gander
University of Geneva, 2-4 rue du Lièvre, Case postale 64, 1211 Genève 4, Switzerland
e-mail: yves.courvoisier@unige.ch

M. J. Gander
e-mail: martin.gander@unige.ch

systems of ordinary differential equations (ODEs) with an iterative process based on a partitioning of the system into subsystems, and solving problems over so called time windows. WR algorithms are also interesting mathematically, since they have often different convergence regimes: for example for linear dissipative systems of ODEs on long time windows, convergence is in general linear, see for example [4, 20, 23, 24], whereas over short time windows, superlinear convergence is often observed for very general systems of linear and non-linear ODEs, see for example [1, 3, 24]. Similarly, for evolution partial differential equations (PDEs), Schwarz waveform relaxation algorithms (SWR) based on an overlapping decomposition of the spatial domain were defined in [16], where also linear convergence was proved over long time windows for the heat equation. The same algorithm also converges superlinearly over short time windows, see [17, 18], and the superlinear convergence rate is faster than for general systems of ODEs, due to the diffusion in the PDE, see also [6] for the more general case of advection reaction diffusion problems. Waveform relaxation algorithms have however often a significant drawback: their convergence can be rather slow. The reason for this lies in the information exchange between subsystems or subdomains, which is performed in the classical WR and SWR through function values or Dirichlet transmission conditions (TCs). In order to obtain more effective algorithms, one has to use optimized TCs, see for example [8, 10] for the case of circuit simulation, where now voltage as well as current values are exchanged between subsystems, and very hard problems can be effectively solved [11]. For PDEs of advection reaction diffusion type, see [2, 9] and references therein. For the wave equation, classical and optimized SWR algorithms have been analyzed in detail in [14, 15], where further interesting convergence behaviors were found (e.g. convergence in a finite number of steps). TCs turned out to be important even for Schwarz methods applied to steady problems, for a review, see [12]. An important further feature of optimized TCs is that the algorithms can then also be used without overlap, in contrast to the classical variants that require overlap for convergence.

In order to optimize TCs in all the WR algorithms applied to diffusive problems described above, it was assumed that the time windows are long, so that the algorithms are in their linear convergence regime. A natural question to ask is how one should optimize the transmission conditions over short time windows, i.e. when the algorithms are in their superlinear convergence regime. It turns out that this analysis is much more difficult, and we present in this paper a first step into this direction for the particular case of the heat equation, and a SWR algorithm with Robin TCs. We first give a very general convergence analysis of the algorithm applied to the d-dimensional heat equation using energy estimates for the case of non-overlapping subdomains. This analysis does however not reveal how the optimized Robin parameter should be chosen. We then focus on the one-dimensional case for which we can give a complete analysis of the algorithm using Laplace transform techniques and explicit estimates of the kernels arising in the iteration. We first obtain an approximation of the optimized parameter on bounded time windows $[0, T]$ by scaling the time domain and applying a frequency analysis of the algorithm, see also [13]. Using the fact that the frequency range is bounded, $\omega \in [\omega_{\min}, \omega_{\max}]$ where one can estimate $\omega_{\min} = \frac{\pi}{T}$ and $\omega_{\max} = \frac{\pi}{\Delta t}$, Δt being the time step of the time discretization, we obtain a first estimate on bounded time windows for the optimized Robin parameter,

which behaves like $\frac{1}{\sqrt{T}}$. Using our explicit expressions of the kernels, and asymptotic analysis, we then obtain a second optimized parameter of the Robin TCs for short time windows, and this parameter behaves like $\frac{1}{T}$. We finally show in a numerical experiment that indeed the optimized parameter changes its asymptotic behavior as the time window becomes small.

2 The Schwarz waveform relaxation algorithm

We consider as our model problem the heat equation in d spatial dimensions,

$$\partial_t u(\mathbf{x}, t) = \Delta u(\mathbf{x}, t) + f(\mathbf{x}, t), \tag{1}$$

where $\mathbf{x} \in \Omega \subset \mathbb{R}^d$ and $f(\mathbf{x}, t)$ is a given source function. This equation is complemented with an initial condition $u_0(\mathbf{x})$, and suitable boundary conditions, for example of Dirichlet type, $u(\mathbf{x}, t) = g(\mathbf{x}, t)$, on the boundary of Ω, denoted by $\Gamma_0 = \partial\Omega$. We first subdivide Ω into non-overlapping subdomains $\tilde{\Omega}_i$, $i \in \{1, 2, \ldots, I\}$, and then enlarge each subdomain $\tilde{\Omega}_i$ by a layer of width at least δ in order to obtain overlapping subdomains Ω_i, see Fig. 1 for a three-dimensional illustration of this decomposition for the example of a cube. For the non-overlapping variant of the algorithm, we simply take $\Omega_i = \tilde{\Omega}_i$, which also corresponds to taking $\delta = 0$ in the overlapping subdomain definition. The boundary of each subdomain Ω_i consists of different parts: if the boundary of Ω_i contains a part of the physical boundary, we denote it by $\Gamma_{i0} := \partial\Omega_i \cap \Gamma_0$. The interior parts of the boundaries of Ω_i are denoted by $\Gamma_{ij} := \partial\Omega_i \cap \overline{\tilde{\Omega}}_j$. The classical Schwarz waveform relaxation algorithm (SWR) applied to the heat equation (1) is then for $n = 1, 2, \ldots$ given by

$$\begin{aligned}
\partial_t u_i^n(\mathbf{x}, t) &= \Delta u_i^n(\mathbf{x}, t) + f(\mathbf{x}, t), && \Omega_i \times (0, T),\\
u_i^n(\mathbf{x}, 0) &= u_0(\mathbf{x}), && (0, T),\\
u_i^n(\mathbf{x}, t) &= g(\mathbf{x}, t), && \Gamma_{i0} \times (0, T),\\
u_i^n(\mathbf{x}, t) &= u_j^{n-1}(\mathbf{x}, t), && \Gamma_{ij} \times (0, T),
\end{aligned} \tag{2}$$

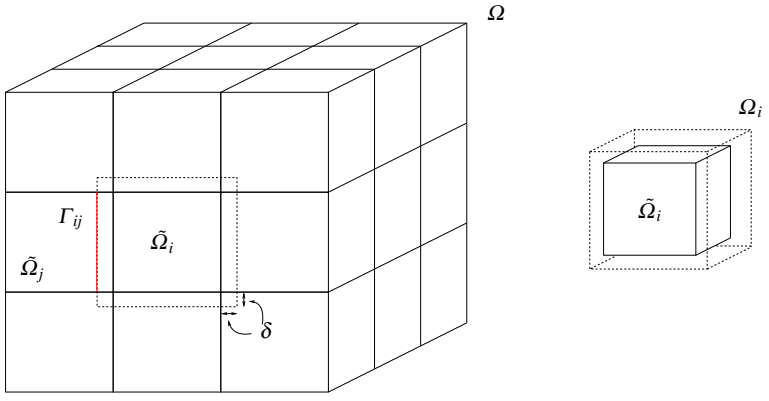

Fig. 1 Decomposition of the domain Ω into overlapping subdomains Ω_i with an overlap of at least 2δ

and one needs an initial guess $u_j^0(\mathbf{x},t)$ to start the algorithm. The last line of (2) is called the *transmission condition* (TC) of the algorithm. Here only Dirichlet information is exchanged on the interfaces, thus we call these TCs the *Dirichlet transmission conditions*. We note that this algorithm does not converge if there is no overlap, i.e. if $\delta = 0$, since the iteration then stagnates on the interfaces: whatever the initial guess on the interface was, it will stay the same, since the Dirichlet trace is taken precisely where the Dirichlet transmission condition was imposed; no new information is thus exchanged, if there is no overlap. More sophisticated transmission conditions are required for non-overlapping subdomains. We focus here on the SWR algorithm with Robin transmission conditions. This algorithm computes for $n = 1, 2, \ldots$ solutions of

$$\begin{aligned}
\partial_t u_i^n(\mathbf{x},t) &= \Delta u_i^n(\mathbf{x},t) + f(\mathbf{x},t), && \Omega_i \times (0,T),\\
u_i^n(\mathbf{x},0) &= u_0(\mathbf{x}), && (0,T),\\
u_i^n(\mathbf{x},t) &= g(\mathbf{x},t), && \Gamma_{i0} \times (0,T),\\
(\partial_{\mathbf{n}_i} u_i^n + p u_i^n)(\mathbf{x},t) &= \left(\partial_{\mathbf{n}_i} u_j^{n-1} + p u_j^{n-1}\right)(\mathbf{x},t), && \Gamma_{ij} \times (0,T),
\end{aligned} \tag{3}$$

where the vector $\mathbf{n}_i$ stands for the unit outward normal vector on the boundary $\partial\Omega_i$. Like before, an initial guess $u_j^0(\mathbf{x},t)$ is needed to start the algorithm. This algorithm converges significantly faster than the classical algorithm (2) for the overlapping case, i.e. for $\delta > 0$, and also converges without overlap, i.e. for $\delta = 0$. The well-posedness of this algorithm with and without overlap in the appropriate Sobolev space setting was studied in [13], and for the more general case of advection reaction diffusion problems in [9]. The convergence analysis of this algorithm, and the optimized choice of the parameter p in it is the focus of this paper. We are interested in both cases, with and without overlap, and the types of analysis one can use depends on this.

3 Convergence analysis with energy estimates

We present in this section a general convergence analysis of algorithm (3) using energy estimate techniques introduced by Lions [22] and Deprès [7] for steady problems, see also [15] for the wave equation. This very general analysis can only be used for the non-overlapping variant of the algorithm, $\delta = 0$, in which case the interfaces satisfy the symmetry relation $\Gamma_{ij} = \Gamma_{ji}$. By linearity, it suffices to study the homogeneous version of algorithm (3), the so called error equations,

$$\begin{aligned}
\partial_t u_i^n(\mathbf{x},t) &= \Delta u_i^n(\mathbf{x},t), && \text{on } \Omega_i \times (0,T),\\
u_i^n(\mathbf{x},0) &= 0, && \text{on } (0,T),\\
u_i^n(\mathbf{x},t) &= 0, && \text{on } \Gamma_{i0} \times (0,T),\\
(\partial_{\mathbf{n}_i} + p) u_i^n(\mathbf{x},t) &= (\partial_{\mathbf{n}_i} + p) u_j^{n-1}(\mathbf{x},t), && \text{on } \Gamma_{ij} \times (0,T),
\end{aligned} \tag{4}$$

and to prove convergence to zero. We assume in what follows that $p > 0$, and consider the energy

$$E^{i,n}(t) = \frac{1}{2}\int_{\Omega_i} u_i^n(\mathbf{x},t)^2 d\mathbf{x}, \tag{5}$$

for which we have the following energy inequality:

Lemma 1 (Energy inequality) *For u_i^n solution of (4) over Ω_i, the energy $E^{i,n}(t)$ defined in (5) satisfies the energy inequality*

$$\partial_t E^{i,n}(t) - \sum_{\substack{j\in V_i \\ j\neq 0}} \int_{\Gamma_{ij}} u_i^n \partial_{\mathbf{n}_i} u_i^n d\gamma \leq 0,$$

where $V_i := \{j \mid \partial\Omega_j \cap \overline{\Omega}_i \neq \emptyset\}$ and $0 \in V_i$ if Γ_{i0} is non empty.

Proof Taking a time derivative of the energy and using the heat equation, we obtain

$$\partial_t E^{i,n}(t) = \int_{\Omega_i} u_i^n \partial_t u_i^n d\mathbf{x} = \int_{\Omega_i} u_i^n \Delta u_i^n d\mathbf{x}.$$

Integrating by parts and dropping the negative term, we obtain the inequality

$$\partial_t E^{i,n}(t) = -\int_{\Omega_i} |\nabla u_i^n|^2 d\mathbf{x} + \int_{\partial\Omega_i} u_i^n \partial_{\mathbf{n}_i} u_i^n d\gamma \leq \sum_{j\in V_i} \int_{\Gamma_{ij}} u_i^n \partial_{\mathbf{n}_i} u_i^n d\gamma.$$

The homogeneous boundary condition in (4) then allows us to eliminate the physical part of the boundary $\partial\Omega_i$ in the sum, which concludes the proof. □

With the notation $\mathcal{T}_p^{i,+} := \partial_{\mathbf{n}_i} + p$ and $T_p^{i,-} := \partial_{\mathbf{n}_i} - p$, we can rewrite the transmission conditions as

$$\mathcal{T}_p^{i,+}\left(u_i^n\right) = \mathcal{T}_p^{i,+}\left(u_j^{n-1}\right), \quad \text{on } \Gamma_{ij}, \tag{6}$$

and computing $\mathcal{T}_p^{i,\pm}(u_i^n)^2 = (\partial_{\mathbf{n}_i} u_i^n)^2 + p^2(u_i^n)^2 \pm 2pu_i^n \partial_{\mathbf{n}_i} u_i^n$, we can reformulate the energy inequality with components that have a sign,

$$\partial_t E^{i,n}(t) - \frac{1}{4p} \sum_{\substack{j\in V_i \\ j\neq 0}} \int_{\Gamma_{ij}} \left(\mathcal{T}_p^{i,+}\left(u_i^n\right)^2 - \mathcal{T}_p^{i,-}\left(u_i^n\right)^2\right) d\gamma \leq 0, \tag{7}$$

which leads to the following convergence result.

Theorem 1 (Convergence in energy) *The Schwarz Waveform Relaxation algorithm (3) for the d-dimensional heat equation converges in the energy norm, i.e.*

$$\sum_{i\in I} E^{i,n}(T) \longrightarrow 0, \quad \text{when } n \longrightarrow +\infty,$$

with I the set of indices of all subdomains and $E^{i,n}$ the energy defined in (5).

Proof We consider the energy inequality (7) and use the transmission condition recalled in (6) to obtain the inequality

$$\partial_t E^{i,n}(t) - \frac{1}{4p} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \left(\mathcal{T}_p^{i,+} \left(u_j^{n-1} \right)^2 - \mathcal{T}_p^{i,-} \left(u_i^n \right)^2 \right) d\gamma \leq 0.$$

Summing over all subdomains, we obtain

$$\partial_t \sum_{i \in I} E^{i,n}(t) - \frac{1}{4p} \sum_{i \in I} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \left(\mathcal{T}_p^{i,+} \left(u_j^{n-1} \right)^2 - \mathcal{T}_p^{i,-} \left(u_i^n \right)^2 \right) d\gamma \leq 0.$$

The key argument now is that we can transform $\mathcal{T}_p^{i,+}(u_j^{n-1})^2$ into $\mathcal{T}_p^{j,-}(u_j^{n-1})^2$ by noting that $\partial_{\mathbf{n}_i} = -\partial_{\mathbf{n}_j}$ on Γ_{ij} ($\mathbf{n}_i = -\mathbf{n}_j$), which implies

$$\begin{aligned} \mathcal{T}_p^{i,+} \left(u_j^{n-1} \right)^2 &= \left(\left(\partial_{\mathbf{n}_i} + p \right) u_j^{n-1} \right)^2 \\ &= \left((-\partial_{\mathbf{n}_j} + p) u_j^{n-1} \right)^2 \\ &= \left((\partial_{\mathbf{n}_j} - p) u_j^{n-1} \right)^2 \\ &= \mathcal{T}_p^{j,-} \left(u_j^{n-1} \right)^2. \end{aligned}$$

We therefore obtain the inequality

$$\begin{aligned} \partial_t \sum_{i \in I} E^{i,n}(t) \; &- \; \frac{1}{4p} \sum_{i \in I} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \mathcal{T}_p^{j,-} \left(u_j^{n-1} \right)^2 d\gamma \\ &+ \; \frac{1}{4p} \sum_{i \in I} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \mathcal{T}_p^{i,-} \left(u_i^n \right)^2 d\gamma \leq 0. \end{aligned}$$

In order to simplify this inequality further, we introduce the quantities

$$\hat{E}^n(t) := \sum_{i \in I} E^{i,n}(t), \quad \hat{\mathcal{T}}^n := \frac{1}{4p} \sum_{i \in I} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \mathcal{T}_p^{j,-} \left(u_j^n \right)^2 d\gamma.$$

We can equivalently switch the dummy summation indices i and j in $\hat{\mathcal{T}}^n$, and write

$$\hat{\mathcal{T}}^n = \frac{1}{4p} \sum_{j \in I} \sum_{\substack{i \in V_j \\ i \neq 0}} \int_{\Gamma_{ji}} \mathcal{T}_p^{i,-} \left(u_i^n \right)^2 d\gamma.$$

Now observe that in this latter formula for $\hat{\mathcal{T}}^n$, summing over all subdomains Ω_j and all interfaces Γ_{ji} for $i \in V_j$, $i \neq 0$ is equivalent to summing over all subdomains

Ω_i and all interfaces Γ_{ij} for $j \in V_i$, $j \neq 0$, since we simply sum over all interfaces $\Gamma_{ij} = \Gamma_{ji}$ twice and the order is not important. We therefore obtain

$$\hat{\mathcal{T}}^n = \frac{1}{4p} \sum_{j \in I} \sum_{\substack{i \in V_j \\ i \neq 0}} \int_{\Gamma_{ji}} \mathcal{T}_p^{i,-} \left(u_i^n\right)^2 d\gamma = \frac{1}{4p} \sum_{i \in I} \sum_{\substack{j \in V_i \\ j \neq 0}} \int_{\Gamma_{ij}} \mathcal{T}_p^{i,-} \left(u_i^n\right)^2 d\gamma.$$

The energy inequality can therefore be written in the more compact form

$$\partial_t \hat{E}^n(t) - \hat{\mathcal{T}}^{n-1} + \hat{\mathcal{T}}^n \leq 0.$$

This suggests to sum over the iterates, leading to a telescopic sum,

$$\partial_t \sum_{n=1}^{N} \hat{E}^n(t) - \hat{\mathcal{T}}^0 + \hat{\mathcal{T}}^N \leq 0,$$

where $\hat{\mathcal{T}}^N$ is positive and thus can be dropped from the inequality, and we obtain

$$\partial_t \sum_{n=1}^{N} \hat{E}^n(t) \leq \hat{\mathcal{T}}^0.$$

Integrating in time and using the fact that the initial condition of the error equations (4) is zero, we obtain

$$\sum_{n=1}^{N} \hat{E}^n(T) \leq \int_0^T \hat{\mathcal{T}}^0 dt.$$

The right hand side is certainly finite and does not depend on the iterate n, hence the left hand side converges when N goes to infinity and thus, since $\hat{E}^n(T) \geq 0$, the principal term of the sum must tend to zero which means that $\hat{E}^n(T)$ tends to zero when the iteration number goes to infinity. □

Although this result ensures convergence of the optimized SWR algorithm for a very general situation, it unfortunately does not give any information about the convergence rate, and the influence of the parameter p on it. We present a much more refined analysis in the next section, albeit for a more simplified situation.

4 Optimization of the algorithm

In order to get more precise convergence estimates for the optimized SWR algorithm with Robin transmission conditions, we focus from now on on the one-dimensional heat equation over the entire real axis, and are interested in bounded solutions u. We decompose the real axis into two subdomains only,

$$\Omega_1 = (-\infty, L), \quad \Omega_2 = (0, +\infty), \quad L \geq 0,$$

which overlap for $L > 0$ and do not overlap for $L = 0$; we thus have both variants of the algorithm we are interested in. The optimized SWR algorithm is then for $n = 1, 2, \ldots$ given by

$$\begin{aligned}
\partial_t u_1^n(x,t) &= \partial_{xx} u_1^n(x,t) + f(x,t), && \Omega_1 \times (0,T),\\
(\partial_x + p) u_1^n(L,t) &= (\partial_x + p) u_2^{n-1}(L,t), && (0,T),\\
u_1^n(x,0) &= u_0(x), && \Omega_1,\\
\partial_t u_2^n(x,t) &= \partial_{xx} u_2^n(x,t) + f(x,t), && \Omega_2 \times (0,T),\\
(\partial_x - p) u_2^n(0,t) &= (\partial_x - p) u_1^{n-1}(0,t), && (0,T),\\
u_2^n(x,0) &= u_0(x), && \Omega_2.
\end{aligned} \tag{8}$$

We define the error $e_i^n := u - u_i^n$, and by linearity, it suffices to study for $n = 1, 2, \ldots$ the associated error equations

$$\begin{aligned}
\partial_t e_i^n(x,t) &= \partial_{xx} e_i^n(x,t), && \Omega_i \times (0,T),\\
\left(\partial_x + (-1)^{i-1} p\right) e_i^n(\Gamma_i,t) &= \left(\partial_x + (-1)^{i-1} p\right) e_j^{n-1}(\Gamma_i,t), && (0,T),\ j \neq i,\\
e_i^n(x,0) &= 0, && \Omega_i,
\end{aligned} \tag{9}$$

where $\Gamma_1 = L$ and $\Gamma_2 = 0$ are the interfaces, i.e. the boundaries of Ω_1 and Ω_2. Like the solutions u and u_i^n, the errors e_i^n must also stay bounded when $|x|$ goes to infinity.

4.1 Laplace analysis

We apply a Laplace transform in time to the error equations (9), where the Laplace transform is defined by

$$\hat{e}_i^n(x,s) = \int_0^\infty e_i^n(x,t) e^{-ts} dt, \quad \Re(s) > 0.$$

This transforms the error equations (9) into a system of ODEs, namely

$$\begin{aligned}
s\hat{e}_i^n(x,s) &= \partial_{xx} \hat{e}_i^n(x,s),\\
\left(\partial_x + (-1)^{i-1} p\right) \hat{e}_i^n(\Gamma_i,s) &= \left(\partial_x + (-1)^{i-1} p\right) \hat{e}_j^{n-1}(\Gamma_i,s),\ j \neq i.
\end{aligned} \tag{10}$$

The Laplace variable s lies in the complex plane with positive real part and the solution is required to stay bounded when $x \in \Omega_i$ tends to infinity. System (10) has the general solution

$$\hat{e}_1^n(x,s) = c_{1,1} e^{-\sqrt{s}(x-L)} + c_{1,2} e^{\sqrt{s}(x-L)}, \quad \hat{e}_2^n(x,s) = c_{2,1} e^{-\sqrt{s}x} + c_{2,2} e^{\sqrt{s}x},$$

where $c_{i,j}$, $i, j = 1, 2$ are functions of s only. The growing exponential term must vanish, since we require that the solutions stay bounded in the spatial variable. Therefore, the solutions simplify to $\hat{e}_1^n(x,s) = c_{1,2} e^{\sqrt{s}(x-L)}$ and $\hat{e}_2^n(x,s) = c_{2,1} e^{-\sqrt{s}x}$, where $c_{1,2}$ and $c_{2,1}$ are determined by the transmission conditions. Computing the derivatives $\partial_x \hat{e}_1^n = \sqrt{s}\hat{e}_1^n$ and $\partial_x \hat{e}_2^n = -\sqrt{s}\hat{e}_2^n$, we obtain the transmission conditions in the frequency domain, for example for the first subdomain

$$(p + \sqrt{s})\hat{e}_1^n(L,s) = (p - \sqrt{s})\hat{e}_2^{n-1}(L,s).$$

This gives for the error $\hat{e}_1^n(x,s)$ on the first subdomain the formula

$$\hat{e}_1^n(x,s) = e^{\sqrt{s}(x-L)}\frac{p-\sqrt{s}}{p+\sqrt{s}}\hat{e}_2^{n-1}(L,s).$$

Similarly, we obtain for the error on the second subdomain

$$\hat{e}_2^n(x,s) = e^{-\sqrt{s}x}\frac{p-\sqrt{s}}{p+\sqrt{s}}\hat{e}_1^{n-1}(0,s).$$

By induction, we obtain after $2n$ iterations

$$\hat{e}_i^{2n}(\Gamma_j,s) = e^{-2nL\sqrt{s}}\left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)^{2n} e_i^0(\Gamma_j,s), \quad i,j=1,2,\ j\neq i.$$

The solution of (9) is the inverse Laplace transform of $\hat{e}_i^{2n}(x,s)$. We must therefore study

$$e_i^{2n}(x,t) = \mathcal{L}^{-1}\left(e^{-2nL\sqrt{s}}\left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)^{2n}\hat{e}_i^0(\Gamma_j,s)\right),\ \ i\neq j.$$

To do so, we compute the inverse Laplace transform of the first factor,

$$f_{2n}(L,t) := \mathcal{L}^{-1}\left(e^{-2nL\sqrt{s}}\left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)^{2n}\right), \tag{11}$$

such that the error is then the convolution

$$e_i^{2n}(\Gamma_j,t) = \int_0^t f_{2n}(L,t-\tau)e_i^0(\Gamma_j,\tau)d\tau. \tag{12}$$

To simplify the notation, we will always consider in what follows n iterations, $n = 1,2,\ldots$, instead of $2n$.

4.2 Explicit Kernel formulas

First we prove technical lemmas leading to a closed form expression for f_n for every positive integer $n = 1,2,\ldots$ and for $L > 0$. We define the functions

$$h_1(y) := \frac{e^{-py}}{\sqrt{py}} + \sqrt{p}\,\mathrm{erf}(\sqrt{py}),$$

$$h_{2,n}(y) := \frac{2^n n! e^{-py}}{(2n)!\sqrt{\pi y}}\mathrm{He}_{2n}(2\sqrt{py}), \tag{13}$$

where He_{2n} is the Hermite polynomial of degree $2n$ defined by the relation,

$$\mathrm{He}_n(y) = (-1)^n e^{y^2/2}\frac{d^n}{dy^n}e^{-y^2/2}. \tag{14}$$

Lemma 2 *For the convolution $h(x) := \int_0^x h_1(x-y)h_{2,n}(y)dy$, we have that*

$$h(0) = \frac{2^n n!\sqrt{\pi}}{(2n)!\sqrt{p}}, \quad n = 1,2,\ldots.$$

Proof The product $h_1(x-y)h_{2,n}(y)$ can be expanded as

$$\begin{aligned} &h_1(x-y)h_{2,n}(y) \\ &= \left(\frac{e^{-p(x-y)}}{\sqrt{p(x-y)}} + \sqrt{p}\operatorname{erf}(\sqrt{p(x-y)})\right)\frac{2^n n! e^{-py}}{(2n)!\sqrt{\pi y}}\operatorname{He}_{2n}(2\sqrt{py}) \\ &= \frac{2^n n!}{(2n)!\sqrt{\pi}}\left(\frac{e^{-px}}{\sqrt{p(x-y)y}} + \frac{\sqrt{p}\operatorname{erf}(\sqrt{p(x-y)})e^{-py}He_{2n}(2\sqrt{py})}{\sqrt{y}}\right). \end{aligned}$$

Thus, we have that $h(x)$ is equal to

$$\begin{aligned} h(x) &= \frac{2^n n!}{(2n)!\sqrt{\pi}}\int_0^x \left(\frac{e^{-px}}{\sqrt{p(x-y)y}}\right. \\ &\quad \left. + \frac{\sqrt{p}\operatorname{erf}(\sqrt{p(x-y)})e^{-py}He_{2n}(2\sqrt{py})}{\sqrt{y}}\right)dx \\ &= \frac{2^n n!}{(2n)!\sqrt{\pi}}\left(\frac{e^{-px}}{\sqrt{p}}\int_0^x \frac{1}{\sqrt{(x-y)y}}dx\right. \\ &\quad \left. + \sqrt{p}\int_0^x \frac{\operatorname{erf}(\sqrt{p(x-y)})e^{-py}He_{2n}(2\sqrt{py})}{\sqrt{y}}dx\right). \end{aligned}$$

We observe then that

$$0 \le \int_0^x \frac{\operatorname{erf}(\sqrt{p(x-y)})e^{-py}He_{2n}(2\sqrt{py})}{\sqrt{y}}dx \le \int_0^x \frac{He_{2n}(2\sqrt{py})}{\sqrt{y}}dx,$$

and that when x goes to zero we have

$$\lim_{x\to 0}\int_0^x \frac{He_{2n}(2\sqrt{py})}{\sqrt{y}}dx = 0,$$

since the function $\frac{He_{2n}(2\sqrt{py})}{\sqrt{y}}$ is integrable. On the other hand, we have that

$$\int_0^x \frac{1}{\sqrt{(x-y)y}}dx = \int_0^1 \frac{1}{(1-t)^{\frac{1}{2}}t^{\frac{1}{2}}}dt \tag{15}$$

with the change of variables $t = \frac{y}{x}$. The integral (15) is given in [19] page 346,

$$\int_0^1 \frac{1}{(1-t)^{\frac{1}{2}}t^{\frac{1}{2}}}dt = \pi,$$

which does not depend on x. We thus obtain

$$h(0) = \lim_{x\to 0} h(x) = \frac{2^n n!\pi}{(2n)!\sqrt{\pi}\sqrt{p}} = \frac{2^n n!\sqrt{\pi}}{(2n)!\sqrt{p}}.$$

□

Lemma 3 *For $h(x) = \int_0^x h_1(x-y)h_{2,n}(y)dy$ where h_1 and $h_{2,n}$ are defined in (13) and $\delta_{nL}(x) = \delta(x - nL)$ the shifted Dirac delta function, we have for $n = 1, 2, \ldots$*

$$\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)e^{-xs}dydx = \int_0^\infty \int_y^\infty h(x-y)\delta'_{nL}(y)e^{-xs}dxdy, \quad (16)$$

where $\delta'_{nL}(y)$ is the weak derivative of the Dirac delta function with respect to y.

Proof We first use integration by parts[1] for the left hand side of (16),

$$\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)e^{-xs}dydx = \int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dye^{-xs}dx$$
$$= \int_0^\infty \left(h(x-y)\delta_{nL}(y)\Big|_0^x - \int_0^x \frac{d}{dy}(h(x-y))\delta_{nL}(y)dy \right) e^{-xs}dx. \quad (17)$$

Because $nL > 0$, we have $\delta_{nL}(0) = 0$, and together with $\frac{d}{dy}(h(x-y)) = -h'(x-y)$, we can simplify (17) to

$$\int_0^\infty \left(h(x-y)\delta_{nL}(y)\Big|_0^x - \int_0^x \frac{d}{dy}(h(x-y))\delta_{nL}(y)dy \right) e^{-xs}dx$$
$$= \int_0^\infty h(0)e^{-xs}\delta_{nL}(x)dx + \int_0^\infty \int_0^x h'(x-y)\delta_{nL}(y)dye^{-xs}dx$$
$$= h(0)e^{-nLs} + \int_0^\infty \int_0^x h'(x-y)\delta_{nL}(y)dye^{-xs}dx$$
$$= h(0)e^{-nLs} + \int_{nL}^\infty h'(x-nL)e^{-xs}dx.$$

Note that $h(0)$ is finite, as shown in Lemma 2.

For the right hand side of (16), we obtain, using the change of variables $z = x - y$,

$$\int_0^\infty \int_y^\infty h(x-y)\delta'_{nL}(y)e^{-xs}dxdy = \int_0^\infty \int_y^\infty h(x-y)e^{-xs}dx\delta'_{nL}(y)dy$$
$$= \int_0^\infty \int_0^\infty h(z)e^{-(z+y)s}dz\delta'_{nL}(y)dy$$
$$= \int_0^\infty \delta'_{nL}(y)e^{-ys}dy \int_0^\infty h(z)e^{-zs}dz. \quad (18)$$

[1] All these calculations hold in the sense of distributions

The first factor $\int_0^\infty \delta'_{nL}(y)e^{-ys}dy$ is the Laplace transform of δ'_{nL} which is equal to

$$\int_0^\infty \delta'_{nL}(y)e^{-ys}dy = se^{-nLs}.$$

Substituting this result into (18) and using again integration by parts leads to

$$\begin{aligned}\int_0^\infty \delta'_{nL}(y)e^{-ys}dy \int_0^\infty h(z)e^{-zs}dz &= se^{-nLs}\int_0^\infty h(z)e^{-zs}dz \\ &= e^{-nLs}\int_0^\infty h(z)se^{-zs}dz \\ &= e^{-nLs}\left(-h(z)e^{-zs}\Big|_0^\infty + \int_0^\infty h'(z)e^{-zs}dz\right) \\ &= h(0)e^{-nLs} + \int_0^\infty h'(z)e^{-zs}e^{-nLs}dz,\end{aligned}$$

where we used in the last equality the fact that $h(z)$ is bounded and $\Re(s) > 0$. Finally, the change of variables $x = z + nL$ gives

$$\int_0^\infty h'(z)e^{-zs}e^{-nLs}dz = \int_{nL}^\infty h'(x - nL)e^{-xs}dx,$$

which shows that the left and right hand sides of (16) are indeed equal. □

Lemma 4 *The Laplace transform of* $h(x) = \int_0^x h_1(x - y)h_{2,n}(y)dy$ *is for* $n = 1, 2, \ldots$ *given by*

$$\mathcal{L}\left(\int_0^x h_1(x - y)h_{2,n}(y)dy\right) = \left(\frac{p - s}{p + s}\right)^n \frac{1}{s}, \tag{19}$$

where h_1 *and* $h_{2,n}$ *are defined in* (13)*, and* p *is a constant.*

Proof The integral $\int_0^x h_1(x - y)h_{2,n}(y)dy$ represents the convolution $h_1 * h_{2,n}$, and the Laplace transform of a convolution is the product of the Laplace transforms, i.e.

$$\mathcal{L}\left(\int_0^x h_1(x - y)h_{2,n}(y)dy\right) = \mathcal{L}(h_1)\mathcal{L}(h_{2,n}).$$

Both Laplace transforms of h_1 and $h_{2,n}$ are given in [25],

$$\mathcal{L}(h_1) = \frac{\sqrt{s + p}}{s}, \qquad \mathcal{L}(h_{2,n}) = \left(\frac{p - s}{p + s}\right)^n \frac{1}{\sqrt{p + s}},$$

which we can multiply to conclude the proof. □

For every function $g(u) \in L^1(\mathbb{R}^+)$, we have the formula

$$\mathcal{L}(g)(\sqrt{s}) = \mathcal{L}\left(\int_0^\infty \eta_t(u)g(u)du\right), \quad \eta_t(u) := \frac{1}{2\sqrt{\pi}t^{\frac{3}{2}}}ue^{-\frac{u^2}{4t}}, \tag{20}$$

see for example [25], page 210. We are ready now to give a closed form expression for the error in the SWR algorithm with Robin transmission conditions.

Theorem 2 *The error of the SWR algorithm* (8) *solved with Robin transmission conditions is for* $n = 1, 2, \ldots$ *of the form*

$$e_i^n(\Gamma_j, t) = \int_0^t f_n(L, t-\tau) e_i^0(\Gamma_j, \tau) d\tau, \tag{21}$$

where the function f_n *is given by*

$$f_n(L, t) = -\int_0^\infty \frac{e^{\frac{-(x+nL)^2}{4t}}(2t - (x+nL)^2)}{4\sqrt{\pi} t^{\frac{5}{2}}} h(x) dx, \tag{22}$$

with $h(x) = \int_0^x h_1(x-y) h_{2,n}(y) dy$ *and* h_1 *and* $h_{2,n}$ *defined in* (13).

Proof We compute first the Laplace transform of the convolution $\int_0^x h(x-y)\delta'_{nL}(y)dy$,

$$\begin{aligned}\mathcal{L}\left(\int_0^x h(x-y)\delta'_{nL}(y)dy\right) &= \int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dy e^{-xs}dx \\ &= \int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)e^{-xs}dydx,\end{aligned}$$

which is, by Lemma 3, equal to

$$\int_0^\infty \int_y^\infty h(x-y)\delta'_{nL}(y)e^{-xs}dxdy.$$

Using (18), we have that

$$\int_0^\infty \int_y^\infty h(x-y)\delta'_{nL}(y)e^{-xs}dxdy = \mathcal{L}(\delta'_{nL})\mathcal{L}(h),$$

where $\mathcal{L}(\delta'_{nL}) = se^{-nLs}$, and by Lemma 4

$$\mathcal{L}(h) = \mathcal{L}\left(\int_0^x h_1(x-y)h_{2,n}(y)dy\right) = \left(\frac{p-s}{p+s}\right)^n \frac{1}{s}.$$

Thus, we obtain

$$\mathcal{L}\left(\int_0^x h(x-y)\delta'_{nL}(y)dy\right) = \left(\frac{p-s}{p+s}\right)^n e^{-nLs},$$

Now using (20) leads to

$$\mathcal{L}\left(\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dy\eta_t(x)dx\right) = \left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)^n e^{-nL\sqrt{s}},$$

which is equivalent to

$$\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dy\eta_t(x)dx = \mathcal{L}^{-1}\left(\left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)^n e^{-nL\sqrt{s}}\right). \tag{23}$$

We now simplify the left hand side of (23). We integrate by parts the inner integral

$$\begin{aligned}
&\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dy\eta_t(x)dx \\
&\quad = \int_0^\infty \left(h(x-y)\delta_{nL}(y)\Big|_0^x + \int_0^x h'(x-y)\delta_{nL}(y)dy \right) \eta_t(x)dx \\
&\quad = \int_0^\infty \left(h(0)\delta_{nL}(x) + \int_0^x h'(x-y)\delta_{nL}(y)dy \right) \eta_t(x)dx \\
&\quad = \int_0^\infty h(0)\delta_{nL}(x)\eta_t(x)dx + \int_0^\infty \int_0^x h'(x-y)\delta_{nL}(y)dy\eta_t(x)dx \\
&\quad = h(0)\eta_t(nL) + \int_{nL}^\infty \int_0^x h'(x-y)\delta_{nL}(y)dy\eta_t(x)dx, \qquad (24)
\end{aligned}$$

where we used the fact that $\delta_{nL}(y) = 0$ for $0 \le x < nL$. We can further simplify the double integral,

$$\int_{nL}^\infty \int_0^x h'(x-y)\delta_{nL}(y)dy\eta_t(x)dx = \int_{nL}^\infty h'(x-nL)\eta_t(x)dx,$$

which we integrate by parts,

$$\begin{aligned}
\int_{nL}^\infty h'(x-nL)\eta_t(x)dx &= h(x-nL)\eta_t(x)\Big|_{nL}^\infty - \int_{nL}^\infty h(x-nL)\eta'_t(x)dx \\
&= -h(0)\eta_t(nL) - \int_{nL}^\infty h(x-nL)\eta'_t(x)dx, \qquad (25)
\end{aligned}$$

where we used that $\lim_{x\to\infty} \eta_t(x) = 0$ and h is bounded. We substitute the result of (25) into (24) and obtain

$$\int_0^\infty \int_0^x h(x-y)\delta'_{nL}(y)dy\eta_t(x)dx = -\int_{nL}^\infty h(x-nL)\eta'_t(x)dx.$$

A last change of variables $\tilde{x} = x - nL$ finally gives

$$\int_{nL}^\infty h(x-nL)\eta'_t(x)dx = \int_0^\infty h(\tilde{x})\eta'_t(\tilde{x}+nL)d\tilde{x},$$

which is equal to (22) and concludes the proof, since then, using (23) and removing the unnecessary tilde on the x variable,

$$f_n(L,t) = \mathcal{L}^{-1}\left(\left(\frac{p-\sqrt{s}}{p+\sqrt{s}} \right)^n e^{-nL\sqrt{s}} \right) = \int_0^\infty h(x)\eta'_t(x+nL)dx.$$

□

4.3 First optimization on bounded time windows

We use in this section the fact that one can scale the heat equation to obtain a time window of constant length (0, 1). This scaling directly reveals the role the time plays

in the convergence behavior of SWR, and leads to a first optimized choice of the parameter p. We also refer to [13], where a similar result has been obtained by direct Fourier analysis, without time scaling.

We consider again the error equations

$$\begin{aligned} \partial_t e_l^n(x,t) &= \partial_{xx} e_l^n(x,t), && \Omega_l \times (0,T), \\ (\partial_{\mathbf{n}_l} + p) e_l^n(\Gamma_l,t) &= (\partial_{\mathbf{n}_l} + p) e_j^{n-1}(\Gamma_l,t), && (0,T),\ j \neq l, \\ e_l^n(x,0) &= 0, && \Omega_l. \end{aligned} \tag{26}$$

We now perform the change of variables to $\tilde{x}$ and $\tilde{t}$ such that $x = \sqrt{T}\tilde{x}$ and $t = T\tilde{t}$. Since (x,t) lies in $\Omega_l \times (0,T)$, we have that $(\tilde{x},\tilde{t})$ belongs to $\Omega_l^s \times (0,1)$ with Ω_l^s defined by

$$\Omega_l^s := \{x \mid \sqrt{T}x \in \Omega_l\}.$$

We define $\tilde{\Gamma}_l$ by $\Gamma_l = \sqrt{T}\tilde{\Gamma}_l$. The error equations in the scaled variables are

$$\begin{aligned} \partial_t e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}) &= \partial_{xx} e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}), && \Omega_l^s \times (0,1), \\ (\partial_{\mathbf{n}_l} + p) e_l^n(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}) &= (\partial_{\mathbf{n}_l} + p) e_j^{n-1}(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}), && (0,1),\ j \neq l, \\ e_l^n(\sqrt{T}\tilde{x}, 0) &= 0, && \Omega_l^s. \end{aligned}$$

Differentiating with respect to x or with respect to $\tilde{x}$ differs only by a constant, i.e. $\partial_{\tilde{x}} = \sqrt{T}\partial_x$, and similarly for the time $\partial_{\tilde{t}} = T\partial_t$. Therefore, the error equations with the tilde differentiation symbols become

$$\begin{aligned} \tfrac{1}{T}\partial_{\tilde{t}} e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}) &= \tfrac{1}{T}\partial_{\tilde{x}\tilde{x}} e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}), && \Omega_l^s \times (0,1), \\ (\partial_{\mathbf{n}_l} + p) e_l^n(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}) &= (\partial_{\mathbf{n}_l} + p) e_j^{n-1}(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}), && (0,1),\ j \neq l, \\ e_l^n(\sqrt{T}\tilde{x}, 0) &= 0, && \Omega_l^s. \end{aligned}$$

Note that the derivative with respect to $\mathbf{n}_l$ is simply a derivative $\partial_{\mathbf{n}_l} = \pm\partial_x$ where the sign depends on which boundary we are located. Hence, a factor $\sqrt{T}$ appears in the transmission conditions,

$$\begin{aligned} (\partial_{\mathbf{n}_l} + p) e_l^n(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}) &= \left(\frac{1}{\sqrt{T}}\partial_{\tilde{\mathbf{n}}_l} + p\right) e_l^n\left(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}\right) \\ (\partial_{\mathbf{n}_l} + p) e_j^{n-1}(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}) &= \left(\frac{1}{\sqrt{T}}\partial_{\tilde{\mathbf{n}}_l} + p\right) e_j^{n-1}\left(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}\right), \end{aligned}$$

where we use $\partial_{\tilde{\mathbf{n}}_l}$ to denote the derivative with respect to $\tilde{x}$ in the direction $\mathbf{n}$. We multiply on both sides of the transmission conditions by $\sqrt{T}$ and simplify the factor $\frac{1}{T}$ in the error equations, to obtain

$$\begin{aligned} \partial_{\tilde{t}} e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}) &= \partial_{\tilde{x}\tilde{x}} e_l^n(\sqrt{T}\tilde{x}, T\tilde{t}), && \Omega_l^s \times (0,1), \\ (\partial_{\tilde{\mathbf{n}}_l} + \sqrt{T}p) e_l^n(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}) &= (\partial_{\tilde{\mathbf{n}}_l} + \sqrt{T}p) e_j^{n-1}(\sqrt{T}\tilde{\Gamma}_l, T\tilde{t}), && (0,1),\ j \neq l, \\ e_l^n(\sqrt{T}\tilde{x}, 0) &= 0, && \Omega_l^s. \end{aligned}$$

Finally, we introduce the new error functions $\tilde{e}_l^n(\tilde{x},\tilde{t}) := e_l^n(\sqrt{T}\tilde{x}, T\tilde{t})$, which leads to error equations equivalent to (26) on the scaled domains $\Omega_l^s \times (0,1)$,

$$\begin{aligned} \partial_{\tilde{t}}\tilde{e}_l^n(\tilde{x},\tilde{t}) &= \partial_{\tilde{x}\tilde{x}}\tilde{e}_l^n(\tilde{x},\tilde{t}), && \Omega_l^s \times (0,1), \\ (\partial_{\tilde{\mathbf{n}}_l} + \sqrt{T}\,p)\tilde{e}_l^n(\tilde{\Gamma}_l,\tilde{t}) &= (\partial_{\tilde{\mathbf{n}}_l} + \sqrt{T}\,p)\tilde{e}_j^{n-1}(\tilde{\Gamma}_l,\tilde{t}), && (0,1),\ j \neq l, \\ \tilde{e}_l^n(\tilde{x},0) &= 0, && \Omega_l^s. \end{aligned} \tag{27}$$

This new formulation illustrates the relative role of the overlap. The overlap between Ω_1^s and Ω_2^s is of size $\frac{L}{\sqrt{T}}$ and thus the overlap is "big" only with respect to the time window, i.e. "big" means $\frac{L}{\sqrt{T}}$ is big. This shows that the absolute size of L is not important, but only the ratio between L and $\sqrt{T}$ is relevant for the convergence speed.

The other main benefit of the scaled formulation is that the Robin parameter is no longer p, but $\sqrt{T}\,p$, and hence the square root of the time window length appears explicitly. We can therefore instantly obtain information about the optimized parameter p^* for bounded time windows. In [13], a direct Fourier analysis was applied to the unscaled error equations (26), which led to the transformed equations

$$\begin{aligned} i\omega\hat{e}_l^n(x,i\omega) &= \partial_{xx}\hat{e}_l^n(x,i\omega), \\ (\sqrt{i\omega} \pm p)\hat{e}_l^n(\Gamma_l,i\omega) &= (\sqrt{i\omega} \pm p)\hat{e}_j^{n-1}(\Gamma_l,i\omega). \end{aligned}$$

Proceeding as before, the solutions are given by

$$\hat{e}_l^n(\Gamma_j,i\omega) = \frac{\sqrt{i\omega}-p}{\sqrt{i\omega}+p} e^{-L\sqrt{i\omega}} \hat{e}_j^{n-1}(\Gamma_l,i\omega), \quad l \neq j,$$

and iterating twice one obtains a recursion with contraction factor ρ,

$$\hat{e}_l^n(\Gamma_j,i\omega) = \rho(i\omega)\hat{e}_l^{n-2}(\Gamma_j,i\omega), \quad \text{with} \quad \rho(i\omega) = \left(\frac{\sqrt{i\omega}-p}{\sqrt{i\omega}+p}\right)^2 e^{-2L\sqrt{i\omega}}. \tag{28}$$

The optimized choice of the parameter p is thus the one that makes the contraction factor as small as possible over all frequencies ω, which leads to the min–max problem

$$\min_{p>0} \max_{\omega \in I} |\rho(i\omega)|, \quad I \subset \mathbb{R}. \tag{29}$$

In a numerical calculation, the range of frequencies lies in the bounded interval $I = (\omega_{\min}, \omega_{\max})$, where one can estimate $\omega_{\max} = \pi/\Delta t$ and $\omega_{\min} = \pi/T$ where $\Delta t = T/m$ with $m+1$ the number of points discretizing the time window $[0,T]$, see [13] for details.

Theorem 3 (Zeroth order approximation, see [13]) *The min–max problem* (29) *with* $I = (\omega_{\min}, \omega_{\max})$ *has, for* $L = 0$*, the optimized parameter* $p^* = (\omega_{\min}\omega_{\max})^{\frac{1}{4}}$*. For* $L > 0$*, three cases occur:*

1. *for* $0 < L \leq \underline{L} = \sqrt{2}\omega_{\min}^{1/4}\omega_{\max}^{-3/4} + o(\omega_{\max}^{-3/4})$, (29) *has for optimized parameter* p^* *the solution of the nonlinear equation* $|\rho(i\omega_{\min})| = |\rho(i\omega_{\max})|$.

2. *for* $\underline{L} < L \leq \bar{L} = C/\sqrt{\omega_{\min}}$, *with* $C = 0.3401\ldots$, (29) *has for optimized parameter* p^* *the solution of the nonlinear equation* $|\rho(i\omega_{\min})| = |\rho(i\omega_e)|$, *with* $\omega_e = pL^{-1}(1 + \sqrt{1 - (Lp)^2 - 2Lp})$.
3. *for* $L > \bar{L}$, (29) *has for optimized parameter* $p^* = \sqrt{\omega_{\min}}$.

The conclusion is here that for small time windows, the overlap of the scaled system (27) is given by $L/\sqrt{T}$ which is big. For a time T small enough such that the condition of the third item of Theorem 3 is satisfied, i.e. $\frac{L}{\sqrt{T}} > \bar{L}$, the theorem gives an estimate for the optimized parameter, namely $p^* = \sqrt{\omega_{\min}} =: p^*_{\text{scaled}}$ and $\omega_{\min}$ is defined for the time interval $(0, 1)$. Recall that the Robin parameter for the scaled problem is $p_{\text{scaled}} = \sqrt{T}\, p_{\text{non-scaled}}$, hence the optimized parameter for the non-scaled problem is

$$p_1^* := p^*_{\text{non-scaled}} = \frac{\omega_{\min}}{\sqrt{T}} = \frac{\sqrt{\pi}}{\sqrt{T}}. \tag{30}$$

This first estimate for the optimized parameter in the SWR algorithm is based on the linear convergence bound (28), and a truncation of the frequencies to the numerically relevant ones. It does however not take into account that the algorithm converges superlinearly on bounded time intervals, a phenomenon which becomes particularly important as soon as one is interested in short time intervals. We show in the next subsection, where we use the exact error expression derived in Section 4.2, that indeed for short times, there is an asymptotically better choice of p.

4.4 Second optimization on short time windows

We now study the convergence behavior of the optimized SWR algorithm (8) on short time intervals, using the explicit kernel representation of f_n from Section 4.2, and we focus on the case of one iteration, i.e. the case $n = 1$. We therefore need to study in detail

$$e_i^1(\Gamma_j, t) = \int_0^t f_1(L, t-\tau) e_i^0(\Gamma_j) d\tau, \tag{31}$$

where f_1 is given by

$$f_1(L, t) = \mathcal{L}^{-1}\left(\frac{p - \sqrt{s}}{p + \sqrt{s}} e^{-L\sqrt{s}}\right).$$

We first compute

$$g(t) = \mathcal{L}^{-1}\left(\frac{p - s}{p + s} e^{-Ls}\right).$$

This inverse Laplace transform is the convolution of the inverse Laplace transforms of

$$\varphi_1(t) := \mathcal{L}^{-1}\left(\frac{p - s}{p + s}\right) \quad \text{and} \quad \varphi_2(t) := \mathcal{L}^{-1}(e^{-Ls}).$$

We have from [5] that $\varphi_1(t) = 2pe^{-pt} - \delta(t)$, and also that $\varphi_2(t) = \delta_L(t)$. Therefore, if $t < L$, we have

$$g(t) = (\varphi_1 * \delta_L)(t) = 0,$$

since $\delta_L(t)$ is zero for all $t < L$. If $t \geq L$, we obtain

$$g(t) = \int_0^t \varphi_1(t-\tau)\varphi_2(\tau)d\tau = \int_0^t (2pe^{-pt} - \delta(t))\delta_L(\tau)d\tau = 2pe^{-p(t-L)} - \delta(t-L).$$

Now by formula (20) we have

$$f_1(L,t) = \mathcal{L}^{-1}\left(\left(\frac{p-\sqrt{s}}{p+\sqrt{s}}\right)e^{-L\sqrt{s}}\right) = \int_0^\infty \eta_t(u)g(u)du = \int_L^\infty \eta_t(u)g(u)du,$$

where

$$\eta_t(u) = \frac{ue^{-\frac{u^2}{4t}}}{2\sqrt{\pi}t^{\frac{3}{2}}}.$$

This then gives the following integral to which we apply the change of variables $z = u - L$,

$$\int_L^\infty \frac{ue^{-\frac{u^2}{4t}}}{2\sqrt{\pi}t^{\frac{3}{2}}}(2pe^{-p(u-L)} - \delta(u-L))du = \int_0^\infty \frac{(z+L)e^{-\frac{(z+L)^2}{4t}}}{2\sqrt{\pi}t^{\frac{3}{2}}}(2pe^{-pz} - \delta(z))dz.$$

This integral can be evaluated in closed form, namely

$$f_1(L,t) = \frac{e^{-\frac{L^2}{4t}}}{2t^{\frac{3}{2}}\sqrt{\pi}}(4pt - L) - 2p^2e^{p^2t+pL}\operatorname{erfc}\left(\frac{L+2pt}{2\sqrt{t}}\right). \tag{32}$$

We can therefore establish an upper bound on e_1^1 in (31), by estimating

$$\begin{aligned}
|e_1^1(\Gamma_2,t)| &\leq \int_0^t |f_1(L,t-\tau)e_1^0(\Gamma_2,\tau)|d\tau \\
&\leq \int_0^t |f_1(L,t-\tau)|d\tau \max_{\tau\in(0,t)}|e_1^0(\Gamma_2,\tau)| \\
&\leq \int_0^t |f_1(L,t-\tau)|d\tau \max_{\tau\in(0,T)}|e_1^0(\Gamma_2,\tau)| \\
&= \int_0^t |f_1(L,t-\tau)|d\tau ||e_1^0(\Gamma_2,\cdot)||_{L^\infty(0,T)}.
\end{aligned}$$

It therefore remains to estimate

$$\int_0^t |f_1(L,t-\tau)|d\tau = -\int_t^0 |f_1(L,\tau)|d\tau = \int_0^t |f_1(L,\tau)|d\tau \leq \int_0^T |f_1(L,\tau)|d\tau,$$

and we thus need an L^1 estimate for f_1 in order to obtain the bound

$$||e_1^1(\Gamma_2,t)||_{L^\infty(0,T)} \leq ||f_1||_{L^1(0,T)}||e_1^0(\Gamma_2,t)||_{L^\infty(0,T)}.$$

A new way to optimize the SWR algorithm with Robin transmission conditions is therefore to determine the parameter p such that the norm $||f_1||_{L^1(0,T)}$ is minimal, i.e.

$$\min_{p>0} ||f_1||_{L^1(0,T)}. \tag{33}$$

Remark 1 The above argument remains unchanged for $n > 1$, i.e. more than one iteration. The associated optimization problem is however much more involved then, and is currently under investigation.

In order to study the optimization problem (33), we first investigate the zeros of f_1 given in (32) using an asymptotic expansion for t small. If t goes to zero, then the fraction $\frac{L+2pt}{2\sqrt{t}}$ goes to infinity. We thus need to expand the complementary error function $\operatorname{erfc}(x)$ into an asymptotic series for $x \to \infty$,

$$\operatorname{erfc}(x) = e^{-x^2}\left(\frac{1}{\sqrt{\pi}x} + O\left(\frac{1}{x^3}\right)\right).$$

This implies that for $t \to 0$ we have the leading order expansion

$$\begin{aligned} e^{p^2t+pL}\operatorname{erfc}\left(\frac{L+2pt}{2\sqrt{t}}\right) &= e^{p^2t+pL}e^{-\left(\frac{L+2pt}{2\sqrt{t}}\right)^2}\left(\frac{2\sqrt{t}}{\sqrt{\pi}(L+2pt)} + O\left(\sqrt{t^3}\right)\right) \\ &= e^{-\frac{L^2}{4t}}\left(\frac{2\sqrt{t}}{\sqrt{\pi}(L+2pt)} + O\left(\sqrt{t^3}\right)\right). \end{aligned}$$

Hence, the function $f_1(L,t)$ has an asymptotic expansion for $t \to 0$ of the from

$$\begin{aligned} f_1(L,t) &= \frac{e^{-\frac{L^2}{4t}}}{2t^{\frac{3}{2}}\sqrt{\pi}}(4pt - L) - 2p^2e^{p^2t+pL}\operatorname{erfc}\left(\frac{L+2pt}{2\sqrt{t}}\right) \\ &= e^{-\frac{L^2}{4t}}\frac{1}{\sqrt{\pi}}\left(\frac{(4pt-L)}{2t^{\frac{3}{2}}} - \frac{4p^2\sqrt{t}}{L+2pt} + O(t)\right) \end{aligned}$$

We can further neglect the term order $\sqrt{t}$, which leads to

$$f_1 \sim e^{-\frac{L^2}{4t}}\frac{(4pt-L)}{\sqrt{\pi}2t^{\frac{3}{2}}} =: \tilde{f}_1.$$

A zero of f_1 is then located around $\tilde{t}_0 := \frac{L}{4p}$ which is the zero of the asymptotic expansion of f_1. This shows also that we expect only one zero of the function $f_1(L,t)$ in the interval $(0,T)$ for small T. We denote the exact zero of $f_1(L,t)$ in this interval by t_0.

Two cases occur: either $\tilde{t}_0$ is in the interval $(0,T)$, or it is greater than T. We first consider the case where $\tilde{t}_0 > T$, thus the L^1-norm of f_1 in the interval $(0,T)$ can be estimated by

$$||f_1||_{L^1(0,T)} \sim ||\tilde{f}_1||_{L^1(0,T)} = \int_0^T |\tilde{f}_1|dt,$$

where we supposed that $\tilde{f}_1$ does not change sign in the interval $(0, T)$. We can then compute the integral using integration by parts,

$$\begin{aligned}
\int_0^T |\tilde{f}_1| dt &= \left| \int_0^T \tilde{f}_1 dt \right| \\
&= \left| \int_0^T e^{-\frac{L^2}{4t}} \frac{(4pt - L)}{\sqrt{\pi} 2t^{\frac{3}{2}}} dt \right| \\
&= \left| \int_0^T \frac{1}{\sqrt{\pi}} e^{-\frac{L^2}{4t}} \left(-\frac{L}{2t^{\frac{3}{2}}} + \frac{2p}{t^{\frac{1}{2}}} \right) dt \right| \\
&= \left| -\frac{L}{2\sqrt{\pi}} e^{-\frac{L^2}{4T}} \frac{4T^{\frac{1}{2}}}{L^2} + \frac{L}{2\sqrt{\pi}} \int_0^T e^{-\frac{L^2}{4t}} \frac{2}{L^2 t^{\frac{1}{2}}} dt + \int_0^T \frac{1}{\sqrt{\pi}} e^{-\frac{L^2}{4t}} \frac{2p}{t^{\frac{1}{2}}} dt \right| \\
&= \left| -e^{-\frac{L^2}{4T}} \frac{2T^{\frac{1}{2}}}{\sqrt{\pi} L} + \left(\frac{1}{\sqrt{\pi} L} + \frac{2p}{\sqrt{\pi}} \right) \int_0^T e^{-\frac{L^2}{4t}} \frac{1}{t^{\frac{1}{2}}} dt \right|.
\end{aligned}$$

Again using integration by parts we evaluate the integral

$$\int_0^T e^{-\frac{L^2}{4t}} \frac{1}{t^{\frac{1}{2}}} dt = e^{-\frac{L^2}{4T}} \frac{4T^{\frac{3}{2}}}{L^2} - \int_0^T e^{-\frac{L^2}{4t}} \frac{6t^{\frac{1}{2}}}{L^2} dt,$$

which gives

$$\begin{aligned}
\int_0^T |\tilde{f}_1| dt = \Bigg| & e^{-\frac{L^2}{4T}} \frac{1}{\sqrt{\pi}} \left(-\frac{2T^{\frac{1}{2}}}{L} + \left(\frac{1}{L} + 2p \right) \frac{4T^{\frac{3}{2}}}{L^2} \right) \\
& - \frac{1}{\sqrt{\pi}} \left(\frac{1}{L} + 2p \right) \int_0^T e^{-\frac{L^2}{4t}} \frac{6t^{\frac{1}{2}}}{L^2} \Bigg|.
\end{aligned}$$

The last integral is of order $e^{-\frac{L^2}{4T}} O(T^{\frac{5}{2}})$, and therefore we can neglect it, i.e.

$$\begin{aligned}
\int_0^T |\tilde{f}_1| dt &= \left| e^{-\frac{L^2}{4T}} \frac{1}{\sqrt{\pi}} \left(-\frac{2T^{\frac{1}{2}}}{L} + \left(\frac{1}{L} + 2p \right) \frac{4T^{\frac{3}{2}}}{L^2} + O(T^{\frac{5}{2}}) \right) \right| \\
&= \left| e^{-\frac{L^2}{4T}} \left(\frac{8T^{\frac{3}{2}}}{\sqrt{\pi} L^2} \left(\frac{L}{4T} - \frac{1}{2L} - p \right) + O(T^{\frac{5}{2}}) \right) \right|. \qquad (34)
\end{aligned}$$

Returning to the optimization problem (33), we thus have that

$$\min_{p>0} ||f_1||_{L^1(0,T)} \sim \min_{p>0} ||\tilde{f}_1||_{L^1(0,T)},$$

and the optimal p is then the one that makes the factor of the exponential vanish in (34). This result is stated in the following theorem.

Theorem 4 *The minimization problem* (33) *has an asymptotic solution* p_2^* *for short time windows given by*

$$p_2^* = \frac{L}{4T} - \frac{1}{2L}.$$

Proof The proof is simply finding the zero of the first term of the right hand side of (34) which vanishes when $p = \frac{L}{4T} - \frac{1}{2L}$. □

Remark 2 We did not investigate the case $\tilde{t}_0 \in (0, T)$ further, since the numerical experiments confirm a dependence of $p_2^* \sim \frac{L}{4T} - \frac{1}{2L}$, which implies that indeed

$$\tilde{t}_0 = \frac{L}{4p_2^*} = \frac{4L^2T}{4(L^2 - 2T)} > \frac{4L^2T}{4L^2} = T.$$

5 Numerical experiment

A first estimate in Section 4.3 showed that the optimized parameter in the Robin transmission conditions for bounded time windows is given by

$$p_1^* = \frac{\sqrt{\pi}}{\sqrt{T}}, \tag{35}$$

see (30). This prediction was obtained through a frequency analysis and by restricting the possible time frequencies to $(\frac{\pi}{T}, \frac{\pi}{\Delta t})$. For short time windows, a second approach taking into account explicit kernel information was shown in Section 4.4. By solving exactly the error equations of the SWR algorithm, we obtained for one iteration step the asymptotically optimized parameter

$$p_2^* = \frac{L}{4T} - \frac{1}{2L}. \tag{36}$$

We compute now in a numerical implementation of the algorithm explicitly the numerically optimized parameter we denote by p_{num}^*. We show in Fig. 2 that indeed this value shows a change of behavior depending on the length of the time window, as predicted by our analysis. For this simulation, we used a centered finite difference discretization in space, with 100 meshpoints for the domain $\Omega = (-1, 1)$, which gives a spatial mesh size $\Delta x = \frac{2}{99}$. In time we used a backward Euler scheme, with constant time step $\Delta t = 10^{-6}$ for all experiments. We used two subdomains and an overlap of $L = \frac{8}{99}$.

Remark 3 The constant used for the red line in Fig. 2 is not the one of the prediction of (35), which can not be expected to be sharp based on the rough numerical frequency estimates. We show in the plot the line defined by $\frac{\sqrt{\pi}}{1.32\sqrt{T}}$.

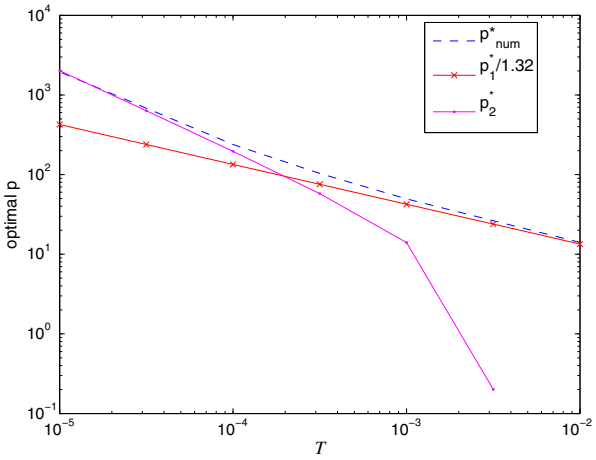

Fig. 2 This figure shows the change of behavior of the numerically computed optimal parameter p^*_{num}. It behaves for short times like the theoretical optimal parameter p^*_2 and for longer times like the theoretical optimal parameter p^*_1 times a constant

6 Conclusions

We have shown in this paper that while the optimized SWR algorithm with Robin transmission conditions can be shown to converge for very general decompositions in many spatial dimensions, the optimization of the Robin parameter over short time windows is a challenging task. This is because the effects of the superlinear convergence regime of SWR need to be taken into account over short time windows. In particular, we have shown that the asymptotic behavior of the optimized parameter necessarily changes from the estimate over bounded time windows, as soon as the time window becomes short, or the overlap large. For the one dimensional heat equation and two subdomains, we provided new explicit kernel formulations for the case of many iterations, and an asymptotic result over short time windows for the optimized parameter over one iteration. Our numerical experiments illustrated the analysis, and also revealed that the optimized parameter based on the simpler linear convergence regime can still be used for quite small time windows.

References

1. Bellen, A., Zennaro, M.: The use of Runge-Kutta formulae in waveform relaxation methods. Appl. Numer. Math. **11**(1–3), 95–114 (1993). Parallel methods for ordinary differential equations (Grado 1991). MR1197152 (94c:65090)
2. Bennequin, D., Gander, M.J., Halpern, L.: A homographic best approximation problem with application to optimized Schwarz waveform relaxation. Math. Comput. **78**(265), 185–232 (2009)
3. Bjørhus, M.: A note on the convergence of discretized dynamic iteration. BIT Numerical Mathematics **35**(2), 291–296 (1995). MR1429020 (97i:65106)
4. Burrage, K.: Parallel and Sequential Methods for Ordinary Differential Equations. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York (1995). Oxford Science Publications. MR1367504 (97f:65021)

5. Colombo, S., Lavoine, J.: Transformations de Laplace et de Mellin. Gauthier-Villars Éditeur, Paris, Formulaires. Mode d'utilisation, Mémorial des Sciences Mathématiques, Fasc. CLXIX. MR0352885 (50 #5371) (1972)
6. Daoud, D.S., Gander, M.J.: Overlapping Schwarz waveform relaxation for advection reaction diffusion problems. Bol. Soc. Esp. Mat. Appl. **46**, 75–90 (2009)
7. Després, B.: Méthodes de décomposition de domaine pour les problèmes de propagation d'ondes en régime harmonique. Le théorème de Borg pour l'équation de Hill vectorielle, Institut National de Recherche en Informatique et en Automatique (INRIA), Rocquencourt, 1991, Thèse, Université de Paris, IX, Dauphine, Paris (1991)
8. Gander, M.J., Al-Khaleel, M., Ruehli, A.: Optimized waveform relaxation methods for longitudinal partitioning of transmission lines. IEEE Trans. Circuits Syst. I **56**(8), 1732–1743 (2009)
9. Gander, M.J., Halpern, L.: Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. SIAM J. Numer. Anal. **45**(2), 666–697 (2007)
10. Gander, M.J., Ruehli, A.: Optimized waveform relaxation methods for RC type circuits. IEEE Trans. Circuits Syst. I **51**(4), 755–767 (2004)
11. Gander, M.J., Ruehli, A.E.: Optimized waveform relaxation solution of electromagnetic and circuit problems. Digest of Electr. Perf. Electronic Packaging (Austin, TX) **19**, 65–68 (2010)
12. Gander, M.J.: Optimized Schwarz methods. SIAM J. Numer. Anal. **44**(2), 699–731 (2006)
13. Gander, M.J., Halpern, L.: Méthodes de relaxation d'ondes (SWR) pour l'équation de la chaleur en dimension 1. C.R. Math. Acad. Sci. Paris **336**(6), 519–524 (2003). MR1975090 (2004a:65124)
14. Gander, M.J., Halpern, L.: Absorbing boundary conditions for the wave equation and parallel computing. Math. Comput. **74**(249), 153–176 (2005). (electronic) MR2085406 (2005h:65158)
15. Gander, M.J., Halpern, L., Nataf, F.: Optimal Schwarz waveform relaxation for the one dimensional wave equation. SIAM J. Numer. Anal. **41**(5), 1643–1681 (2003). MR2035001 (2005h:65252)
16. Gander, M.J., Stuart, A.M.: Space-time continuous analysis of waveform relaxation for the heat equation. SIAM J. Sci. Comput. **19**(6), 2014–2031 (1998). MR1638096 (99h:65164)
17. Gander, M.J., Zhao, H.: Overlapping Schwarz waveform relaxation for the heat equation in n dimensions. BIT Numerical Mathematics **42**(4), 779–795 (2002). MR1944537 (2003m:65170)
18. Giladi, E., Keller, H.B.: Space-time domain decomposition for parabolic problems. Numer. Math. **93**(2), 279–313 (2002). MR1941398 (2003j:65088)
19. Hairer, E., Wanner, G.: L'analyse au fil de l'histoire, Vol. 10. Springer-Verlag, Berlin (2001)
20. Jeltsch, R., Pohl, B.: Waveform relaxation with overlapping splittings. SIAM J. Sci. Comput. **16**(1), 40–49 (1995). MR1311677 (95j:65079)
21. Lelarasmee, E., Ruehli, A.E., Sangiovanni-Vincentelli, A.L.: The waveform relaxation method for time-domain analysis of large scale integrated circuits. IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. CAD-1 **3**, 131–145 (1982)
22. Lions, P.-L.: On the Schwarz alternating method. III. A variant for nonoverlapping subdomains. In: Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989), pp. 202–223. SIAM, Philadelphia, PA (1990)
23. Miekkala, U., Nevanlinna, O.: Convergence of dynamic iteration methods for initial value problem. SIAM J. Sci. Statist. Comput. **8**(4), 459–482 (1987). MR892300 (89f:65076)
24. Nevanlinna, O.: Remarks on Picard-Lindelöf iteration. II. BIT Numerical Mathematics **29**(3), 535–562 (1989). MR1009655 (91b:65088)
25. Oberhettinger, F., Badii, L.: Tables of Laplace Transforms. Springer-Verlag, New York (1973). MR0352889 (50 #5375)