Optimal Control of the Convergence Rate of Schwarz Waveform Relaxation Algorithms

Florian Lemarié[1], Laurent Debreu[2], and Eric Blayo[3]

[1] Institute of Geophysics and Planetary Physics, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024-1567, United States, florian@atmos.ucla.edu

[2] INRIA Grenoble Rhône-Alpes, Montbonnot, 38334 Saint Ismier Cedex, France and Jean Kuntzmann Laboratory, BP 53, 38041 Grenoble Cedex 9, France, laurent.debreu@imag.fr

[3] University of Grenoble and Jean Kuntzmann Laboratory, BP 53, 38041 Grenoble Cedex 9, France, eric.blayo@imag.fr

Summary. In this study we present a *non-overlapping Schwarz waveform relaxation method* applied to the one dimensional unsteady diffusion equation. We derive efficient interface conditions using an *optimal control* approach once the problem is discretized. Those conditions are compared to the usual optimized conditions derived at the PDE level by solving a *min-max problem*. The performance of the proposed methodology is illustrated by numerical experiments.

1 Introduction

Schwarz-like domain decomposition methods are very popular in mathematics, computational sciences, and engineering notably for the implementation of coupling strategies. This type of method, originally introduced for stationary problems, can be extended to evolution problems by adapting the waveform relaxation algorithms to provide the so-called Schwarz waveform relaxation method [2, 4]. The idea behind this method is to separate the spatial domain, over which the time-evolution problem is defined, into subdomains. The resulting time-dependent problems are then solved separately on each subdomains. An iterative process with an exchange of boundary conditions at the interface between the subdomains is then applied to achieve the convergence to the solution of the original problem. To accelerate the convergence speed of the iterative process, it is possible to derive efficient interface conditions by solving an optimization problem related to the convergence rate of the method [e.g.; 1, 5].

In this study, we specifically address the optimization problem arising from the use of *Robin* type transmission conditions in the framework of a *non-overlapping Schwarz waveform relaxation*. For this type of problem, the existing work has been achieved mainly at the *PDE level*, giving rise to the optimized Schwarz waveform

R. Bank et al. (eds.), *Domain Decomposition Methods in Science and Engineering XX*, Lecture Notes in Computational Science and Engineering 91,
DOI 10.1007/978-3-642-35275-1__71,

relaxation algorithm [1, 2, 5]. The objective here is to use the *optimal control theory* paradigm [9] to find parameters optimized at the *discrete level*, and thus to systematically make a comparison with the parameters determined at the PDE level. This paper is organized as follows : in Sect. 2 we briefly recall the basics of optimized Schwarz methods in the framework of a time evolution problem. Section 3 is dedicated to the determination of the *optimal control problem* that we intend to address. Finally, in Sect. 4 we apply our approach to a diffusion problem.

2 Optimization of the Convergence at the PDE Level

2.1 Model Problem and Optimized Schwarz Methods

Let us consider Ω a bounded open set of $\mathbb{R}$. The model problem is to find u such that u satisfies over a time period $[0,T]$

$$\mathscr{L}u = f, \qquad \text{in } \Omega \times [0,T], \tag{1}$$

$$\mathscr{B}u = g, \qquad \text{on } \partial\Omega \times [0,T], \tag{2}$$

where $\mathscr{L}$ and $\mathscr{B}$ are two partial differential operators, and f the forcing. This problem is complemented by an initial condition

$$u(x,0) = u_0(x), \qquad x \in \Omega. \tag{3}$$

We consider a splitting of the domain Ω into two *non-overlapping domains* Ω_1 and Ω_2 communicating through their common interface Γ. The operator $\mathscr{L}$ introduced previously is split into two operators $\mathscr{L}_j$ restricted to Ω_j $(j = 1,2)$. By noting $\mathscr{F}_1$, $\mathscr{F}_2$, $\mathscr{G}_1$ and $\mathscr{G}_2$ the operators defining the interface conditions, the alternating form of the *Schwarz waveform relaxation algorithm* reads

$$\begin{cases} \mathscr{L}_1 u_1^k = f_1, & \text{in } \Omega_1 \times [0,T], \\ u_1^k(x,0) = u_o(x), & x \in \Omega_1, \\ \mathscr{B}_1 u_1^k(x,t) = g_1, & \text{in } [0,T] \times \partial\Omega_1, \\ \mathscr{F}_1 u_1^k(0,t) = \mathscr{F}_2 u_2^{k-1}(0,t), & \text{in } \Gamma \times [0,T], \end{cases} \quad \begin{cases} \mathscr{L}_2 u_2^k = f_2, & \text{in } \Omega_2 \times [0,T], \\ u_2^k(x,0) = u_o(x), & x \in \Omega_2, \\ \mathscr{B}_2 u_2^k(x,t) = g_2, & \text{in } [0,T] \times \partial\Omega_2, \\ \mathscr{G}_2 u_2^k(0,t) = \mathscr{G}_1 u_1^k(0,t), & \text{in } \Gamma \times [0,T], \end{cases} \tag{4}$$

where $k = 1,2,\ldots$ is the iteration number, and the initial guess $u_2^0(0,t)$ must be given. The operators $\mathscr{F}_j$ and $\mathscr{G}_j$ must be chosen to impose the desired consistency of the solution on the interface Γ. We consider here the one-dimensional diffusion equation with constant (possibly discontinuous) diffusion coefficients κ_j $(\kappa_j > 0, j = 1,2)$. We define $\mathscr{L}_j = \partial_t - \kappa_j \partial_x^2$, $\Omega_1 = (-L_1, 0)$, $\Omega_2 = (0, L_2)$ $(L_1, L_2 \in \mathbb{R}^+)$, and $\Gamma = \{x = 0\}$. In this context, we require the equality of the subproblems solutions and of their normal fluxes on the interface Γ,

$$u_1(0,t) = u_2(0,t), \qquad \kappa_1 \partial_x u_1(0,t) = \kappa_2 \partial_x u_2(0,t), \qquad t \in [0,T]. \tag{5}$$

To obtain such a consistency we use mixed boundary conditions of *Robin* type

$$\mathscr{F}_j = -\kappa_j \partial_x + p_1, \qquad \mathscr{G}_j = \kappa_j \partial_x + p_2, \qquad (j = 1, 2),$$

where p_1 and p_2 are two parameters that can be optimally chosen to improve the convergence speed of the Schwarz method. Algorithm (4) with two-sided Robin conditions (i.e. for $p_1 \neq p_2$) is well-posed for any choice of p_1 and p_2 such that $p_1 + p_2 > 0$. This result can be shown using a priori energy estimates, as described in [4].

2.2 Optimization of the Convergence Factor

To demonstrate the convergence of algorithm (4) a classical approach [e.g. 6] is to define the error e_j^k between the exact solution $u^\star$ and the iterates u_j^k. A Fourier analysis enables the transformation of the original PDEs into ODEs that can be solved analytically. The analytical solution on each subdomain is then used to define a convergence factor ρ of the corresponding *Schwarz algorithm*. For a diffusion problem, defined on subdomains of infinite size (i.e. assuming $L_1, L_2 \to \infty$), we get

$$\rho(p_1, p_2, \omega) = \left| \frac{(p_2 - \sqrt{i\omega\kappa_2})}{(p_2 + \sqrt{i\omega\kappa_1})} \frac{(p_1 - \sqrt{i\omega\kappa_1})}{(p_1 + \sqrt{i\omega\kappa_2})} \right|, \tag{6}$$

where p_1 and p_2 are two degrees of freedom which can be tuned to accelerate the convergence speed. In (6), $i = \sqrt{-1}$, and $\omega \in \mathbb{R}$ is the angular frequency arising from a Fourier transform in time on e_j^k. A general approach to choose the Robin parameters p_1 and p_2 is to solve a minimax problem [2]

$$\min_{p_1, p_2 \in \mathscr{R}} \left(\max_{\omega \in [\omega_{\min}, \omega_{\max}]} \rho(p_1, p_2, \omega) \right). \tag{7}$$

Because we work in practice on a discrete problem the frequencies allowed by the temporal grid range from $\omega_{\min} = \pi/T$ to $\omega_{\max} = \pi/\Delta t$, where Δt is the time step of the temporal discretization. For the diffusion problem under consideration here, the analytical solution of the optimization problem (7) has been derived in [8] in a general *two-sided* case (i.e. with $p_1 \neq p_2$) with discontinuous coefficients $\kappa_1 \neq \kappa_2$. For the sake of simplicity, we consider in the present study the continuous case ($\kappa_1 = \kappa_2 = \kappa$) and we recall the result found in [8] in this case.

Theorem 1. *Under the assumption $\kappa_1 = \kappa_2 = \kappa$, the optimal parameters $p_1^\star$ and $p_2^\star$ of the minmax problem (7) are given by*

$$p_1^\star = \frac{\alpha\sqrt{2\kappa}}{4}\left[\sqrt{8+\nu^2} - \nu\right], \qquad p_2^\star = \frac{\alpha\sqrt{2\kappa}}{4}\left[\sqrt{8+\nu^2} + \nu\right],$$

where $\alpha = (\omega_{\min}\omega_{\max})^{1/4}$, $\beta = \alpha^{-1}(\sqrt{\omega_{\min}} + \sqrt{\omega_{\max}})$ and

$$\nu = \begin{cases} 2\sqrt{\beta - 1} & \text{if } \beta \geq 1 + \sqrt{5}, \\ \sqrt{2\beta^2 - 12} & \text{if } \sqrt{6} \leq \beta < 1 + \sqrt{5}, \\ 0 & \text{if } 2 < \beta < \sqrt{6}. \end{cases}$$

It is worth mentioning that even if the diffusion coefficients are continuous the *two-sided* case provides a faster convergence than the *one-sided* case studied in [4] (Fig. 1).

General Remarks :

- The usual methodology to optimize the convergence at the continuous level comes with a few assumptions that may lead to inaccuracies once the problem is discretized. For example, as discussed in [7] (Sect. 5), the *infinite domain assumption* used to determine the convergence factor (6) may lead to appreciable differences in the optimized parameters compared to an approach taking the finiteness of the subdomains into account. We numerically found that the *infinite domain assumption* is valid as long as the *dimensionless Fourier number* $\mathrm{Fo} = \kappa_j/(L_j^2\omega)$ (with L_j the size of subdomain Ω_j) of the problem does not exceed a critical value $\mathrm{Fo}_c = 0.02$.
- The optimization problem (7) aims at minimizing the maximum value of $\rho(p_1, p_2, \omega)$ over the entire interval $[\omega_{\min}, \omega_{\max}]$. This provides a very robust method general enough to deal with the worst case scenario when all the temporal frequencies are present in the error. An even more efficient way to proceed would be to adjust the values of p_1 and p_2 at each iteration so that those parameters are efficiently chosen to "fight" the remaining frequencies in the error.

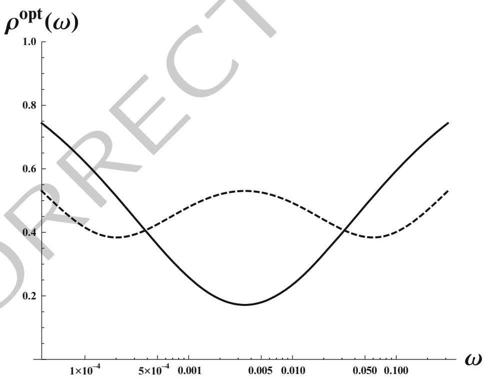

Fig. 1. Convergence factor optimized at the PDE level in the *one-sided* case (*black line*) [4] and in the *two-sided* case (*dashed black line*) [8], for $\kappa = 10^{-2}$ m s^{-1}, $\Delta t = 10$ s, and $T = 2^{13}\Delta t$

3 Optimal Control of the *Robin* Parameters

To investigate the robustness of the optimized parameters once the problem is discretized, the use of the *optimal control theory* appears as a natural choice. We aim at controlling the *Robin* parameter in order to get the best possible convergence speed in the sense of a given cost function $\mathscr{J}$. Moreover, following the approach of [3] and the previous discussion, we consider the possibility to use different parameters p_j for different steps of the iterative process. It is easy to check that by choosing different parameters at each iteration we still converge to the solution of the global problem. A first way to choose the parameters is to look, at each iteration k, for p_1^k and p_2^k minimizing the error at the interface. In this case the cost function that we intend to minimize at each iteration would be

$$\begin{aligned}\mathscr{J}(p_1^k, p_2^k) = &\frac{w}{2}\int_0^T \left(u_1^k(0,t) - u_2^k(0,t)\right)^2 dt \\ &+ \frac{\widetilde{w}}{2}\int_0^T \left(\kappa_1 \partial_x u_1^k(0,t) - \kappa_2 \partial_x u_2^k(0,t)\right)^2 dt.\end{aligned} \tag{8}$$

The constants w and $\widetilde{w}$ must be chosen to balance both terms, depending on the characteristics of the problem (see Sect. 4). The cost function (8) is designed in agreement with the consistency (5) we want to impose at the interface between subdomains. $\mathscr{J}$ provides a measure of the "inconsistency" of the solution at each iteration k, and is, thus, directly related to the order of magnitude of the errors e_j^k of the algorithm (as shown in Fig. 2). An other strategy could be to minimize the error at a given iteration K. The cost function would thus be

$$\begin{aligned}\mathscr{J}\left((p_1^k, p_2^k)_{k=1,K}\right) = &\frac{w}{2}\int_0^T \left(u_1^K(0,t) - u_2^K(0,t)\right)^2 dt \\ &+ \frac{\widetilde{w}}{2}\int_0^T \left(\kappa_1 \partial_x u_1^K(0,t) - \kappa_2 \partial_x u_2^K(0,t)\right)^2 dt,\end{aligned} \tag{9}$$

leading to an optimization on $2K$ parameters. This latter approach is particularly interesting when we intend to obtain the best possible approximation of the exact solution after a number of iterations set in advance. We propose here to lead our study with this kind of approach with $K = 5$. The *optimal control* approach does not *per se* reduce the computational cost of the algorithm because many evaluations of the cost function are required during the minimization process (see Algorithm 3). We use this approach as a tool to improve our understanding of the behavior of the Robin parameters in order to find new directions to further accelerate the convergence speed when Robin-type interface conditions are used. We denote by $\mathbf{p_1^{\star,num}}$ and $\mathbf{p_2^{\star,num}}$ the parameters found numerically by solving the optimal control problem. Those parameters correspond to two vectors of size K. Similarly we will denote by $\mathrm{p}_1^{\star,\mathrm{ana}}$ and $\mathrm{p}_2^{\star,\mathrm{ana}}$ the parameters found analytically (cf. Theorem 1).

We used Matlab for the computation (Algorithm 3). Note that the well-posedness of the coupling problem (4) is not sufficient to ensure a well-posed optimal control problem. Some additional requirements on the convexity and regularity of the cost function are necessary. We do not provide here such a proof, however we empirically checked that the same solution of the optimal problem is obtained for a wide range of parameter values for the initial guess.

4 Numerical Experiments

We discretized problem (4) using a *backward Euler* scheme in time and a second order scheme defined on a staggered grid in space (see [8] for more details). We

Algorithm 3 Optimal control

```
%== Robin parameters found analytically : p1ana, p2ana
%== Solution of the optimal control problem : p1opt, p2opt
%== Initial guess ==%
x0(1:2:2*K-1)=p1ana;
x0(2:2:2*K  )=p2ana;
%== Solve the optimal control problem ==%
%== the CalcJ function proceeds to K iterations of the
%== Schwarz algorithm using 2K Robin parameters,
%== and computes the associated cost function (9)
x = fminsearch( @CalcJ, x0 );
%== Retrieve the optimized parameters
p1opt(1:K)=x(1:2:2*K-1);
p2opt(1:K)=x(2:2:2*K  );
```


decompose the domain Ω into two non-overlapping subdomains $\Omega_1 = [-H,0]$ and $\Omega_2 = [0,H]$ with $H = 500$ m. The diffusion coefficient is $\kappa = 10^{-2}$ m^2 s^{-1} and the total simulation time is $T = 2^{13}\Delta t$ with $\Delta t = 10$ s. The parameter values lead to a *dimensionless Fourier number* smaller than 0.02 so that the *infinite domain assumption* is valid. We simulate directly the error equations, i.e. $f_1 = f_2 = 0$ in (4) and $u_0(x) = 0$. We start the iteration with a random initial guess $u_2^0(0,t)$ $(t \in [0,T])$ so that it contains a wide range of the temporal frequencies that can be resolved by the computational grid. This is done to allow a fair comparison as the parameters optimized at the PDE level are optimized assuming that the full range $[\omega_{\min}, \omega_{\max}]$ is present in the error. We first perform the Optimized non-overlapping Schwarz Method (referred as to OSM case) using $\mathrm{p}_1^{\star,\mathrm{ana}}$ and $\mathrm{p}_2^{\star,\mathrm{ana}}$ and then using an optimal control of the *Robin* parameters with $K = 5$ (referred as to OptCon case). We first check that the minimization of cost function $\mathscr{J}$ consistently implies the reduction of the errors $\|e_j\|_\infty$ of the associated algorithm (Fig. 2). For our experiments, we chose $w = 1$ and $\widetilde{w} = H/\kappa$ in (9). We notice that in the OptCon case the convergence speed is significantly improved compared to the OSM case. Indeed, nine iterations of the OSM are required to obtain the same accuracy than the OptCon case after only five iterations. In order to have more insight on the way the parameters $\mathbf{p}_1^{\star,\mathbf{num}}$ and $\mathbf{p}_2^{\star,\mathbf{num}}$ evolve throughout the iterations we plot, in Fig. 3, the corresponding convergence factor (6) at each iteration. It is striking to realize that the optimal convergence is obtained through a combination of 2-point (equivalent to the *one-sided* case) and 3-point (equivalent to the *two-sided* case) equioscillations sometimes shifted along the ω-axis to adapt to the temporal frequencies still present in the error. The first two iterations aim at working mainly on the high-frequency components while the last three iterations are optimized to work on the low-frequency component. The adaptivity of the *Robin* parameters from one iteration to the other brings more flexibility to the method enabling more scale selectivity.

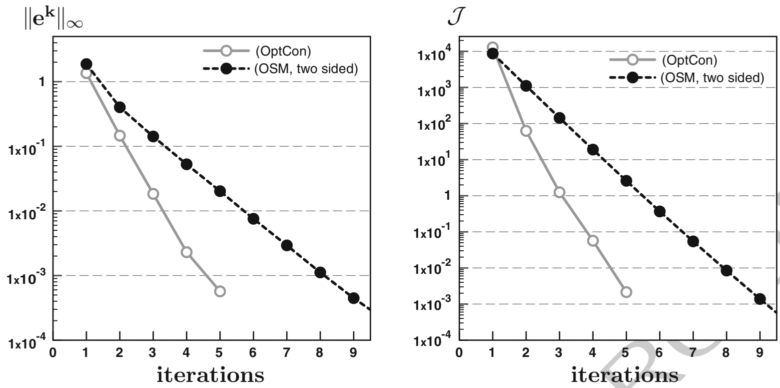

Fig. 2. Evolution of the $\mathscr{L}^\infty$-norm of the error (*left*) and of the cost function $\mathscr{J}$ (*right*) with respect to the iterates k in the OSM and OptCon cases

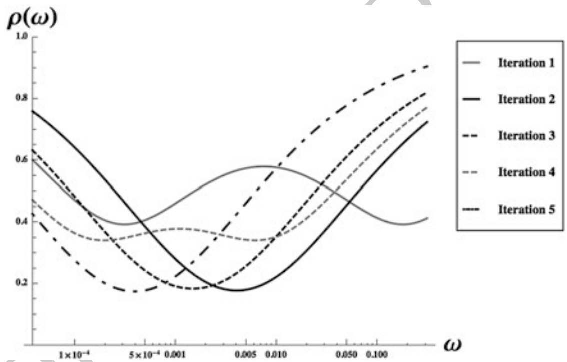

Fig. 3. Sequence of convergence factors $\rho(\omega)$ resulting from the optimal control of the Robin parameters determined to get the best possible convergence after $K = 5$ iterations

5 Conclusion

Due to its simplicity, the use of *Robin-type transmission conditions* is very attractive when one wants to couple unsteady problems defined on non-overlapping subdomains. Once the *Robin* parameters are properly chosen one can achieve a fast convergence [2]. In the present study we showed that there is still room for improvement in the design of the Robin conditions. If the *Robin* parameters are adjusted from one iteration to the other we showed, thanks to an optimal control approach, that we can significantly improve the convergence speed. It is important to emphasize that the *optimal control* paradigm proposed in this study is general enough to be used with any type of PDE and an arbitrary number of subdomains.

Acknowledgments This research was partially supported by the ANR project COMMA (COupling in Multi-physics and multi-scale problems: Models and Algorithms) and by the INRIA project-team MOISE (Modelling, Observation and Identification for Environmental Sciences). We are thankful to Héloïse Pelen (ENS Lyon) for her contribution during her masters internship.

Bibliography

[1] D. Bennequin, M. J. Gander, and L. Halpern. Optimized Schwarz waveform relaxation methods for convection reaction diffusion problems. *Technical Report 2004-24, LAGA, Université Paris 13*, 2004.

[2] M. J. Gander, L. Halpern, and F. Nataf. Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation. In *Eleventh International Conference on Domain Decomposition Methods (London, 1998)*, pages 27–36 (electronic). DDM.org, Augsburg, 1999.

[3] M. J. Gander and G. H. Golub. A non-overlapping optimized Schwarz method which converges with arbitrarily weak dependence on h. In *Domain decomposition methods in science and engineering*, pages 281–288 (electronic). Natl. Auton. Univ. Mex., México, 2003.

[4] M. J. Gander and L. Halpern. Methodes de relaxation d'ondes pour l'equation de la chaleur en dimension 1. *C. R. Acad. Sci. Paris*, 336(Série I):519–524, 2003.

[5] M. J. Gander and L. Halpern. Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems. *SIAM J. Numer. Anal.*, 45(2):666–697 (electronic), 2007. ISSN 0036–1429. doi: 10.1137/050642137.

[6] M. J. Gander, L. Halpern, and M. Kern. A Schwarz waveform relaxation method for advection-diffusion-reaction problems with discontinuous coefficients and non-matching grids. in *Domain decomposition methods in science and engineering XVI*, vol. 55 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, pp. 283–290, 2007.

[7] F. Lemarié, L. Debreu, and E. Blayo. Optimized global-in-time Schwarz algorithm for diffusion equations with discontinuous and spatially variable coefficients. Research Report RR-6663, INRIA, 2008.

[8] F. Lemarié, L. Debreu, and E. Blayo. Toward an optimized global-in-time Schwarz algorithm for diffusion equations with discontinuous and spatially variable coefficients, part 1 : the constant coefficients case. *Electron. Trans. Numer. Anal.*, 2012. (in revision).

[9] J.-L. Lions. *Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles*. Dunod, Paris, 1968.