



**HAL**  
open science

# Coupling Parareal with Optimized Schwarz waveform relaxation for parabolic problems

Duc Quang Bui, Caroline Japhet, Yvon Maday, Pascal Omnes

► **To cite this version:**

Duc Quang Bui, Caroline Japhet, Yvon Maday, Pascal Omnes. Coupling Parareal with Optimized Schwarz waveform relaxation for parabolic problems. 2021. hal-03167571v1

**HAL Id: hal-03167571**

**<https://inria.hal.science/hal-03167571v1>**

Preprint submitted on 12 Mar 2021 (v1), last revised 5 Aug 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Coupling Parareal with Optimized Schwarz waveform relaxation for parabolic problems\*

Duc Quang Bui<sup>‡</sup>, Caroline Japhet<sup>‡</sup>, Yvon Maday<sup>§</sup> and Pascal Omnes<sup>||‡</sup>

## Abstract

We propose and analyse a parallel method, both in the time and space directions, that couples the Parareal algorithm with the Optimized Schwarz waveform relaxation (OSWR) method, with only few OSWR iterations in the fine propagator and with a simple coarse propagator deduced from the Backward Euler method. The analysis of this coupled method is presented for a one-dimensional advection-reaction-diffusion equation. We prove a general convergence result for this method via energy estimates. Numerical results for two-dimensional advection-diffusion problems and for a diffusion equation with strong heterogeneities are presented, to illustrate the performance of the coupled Parareal-OSWR algorithm.

## 1 Introduction

The Parareal algorithm is a time-parallel method that was proposed by Lions, Maday, and Turinici to solve evolutionary problems in parallel, see [23]. The algorithm is constructed using two solvers: the coarse solver, which is fast but not very accurate, and the fine solver, which is slower but more accurate. The long time interval is divided into smaller windows, the fine solver is performed on each window, using some input initial conditions. The outputs are then corrected by the coarse solver and used as the inputs for the next iteration. In the most simplified convergence analysis, the initial value problem was considered to be an ODE, the fine solver was exact and the coarse solver was the Backward Euler method; in that analysis the outputs were computed exactly and so was the error. More complex analyses of the performance of Parareal were also achieved: for non-exact fine solvers for ODEs [3], for linear ordinary and partial differential equations [14]; for the non-linear case [7]; for improvement of the parallel efficiency [27].

In the construction of Parareal, one may choose suitable coarse and fine solvers to fasten the process, for example by using an iterative method like, e.g., the Schwarz waveform relaxation (SWR) method [13]. As Parareal itself is also an iterative method, the overall process would then be composed of outer iterations (Parareal) and inner iterations corresponding to the coarse and fine solvers. In order to save CPU resources, we might think of stopping the inner iterative solvers after a small number of iterations, well before convergence, and hope that the overall convergence could be carried through Parareal iterations. This idea was first raised in [28] for general iterative methods; it was then developed for some iterative methods: for the Jacobi method (see [30]), for SWR methods (see [16], [5]). In this contribution, we are more interested in the latter, which will be recalled next.

SWR algorithms are based on a spatial domain decomposition. The spatial domain is decomposed into overlapping or non overlapping subdomains; then the original problem is transformed into multi-domain space-time sub-problems. Starting from initial fluxes on the space-time interfaces, each sub-problem can then be solved in parallel over the whole time range; then the data are exchanged through the interfaces to

---

\*This work was supported by the ANR project CINE-PARA under grant ANR-15-CE23-0019.

<sup>‡</sup>CNRS, UMR 7539, Laboratoire de Géométrie, Analyse et Applications, LAGA, Université Sorbonne Paris Nord, F-93430, Villetaneuse, France, [bui@math.univ-paris13.fr](mailto:bui@math.univ-paris13.fr), [japhet@math.univ-paris13.fr](mailto:japhet@math.univ-paris13.fr).

<sup>§</sup>Sorbonne Université, Université Paris-Diderot SPC, CNRS, Laboratoire Jacques-Louis Lions (LJLL), F-75005 Paris, France, and Institut Universitaire de France, 75005 Paris, France, [maday@ann.jussieu.fr](mailto:maday@ann.jussieu.fr).

<sup>||</sup>CEA Saclay, DM2S-STMF, F91191 Gif sur Yvette Cedex, France, [pascal.omnes@cea.fr](mailto:pascal.omnes@cea.fr).

create better fluxes. The transmission conditions for the fluxes play an important role in the convergence process and several possibilities can be used, e.g., Dirichlet [15], Robin [29], [8] or Ventcell [22]. The last two types of conditions can be optimized [21, 6]. While in the classical Schwarz method, spatial fluxes are exchanged at each time step [6], the space-time fluxes in the SWR method leave a wide range of choices for the discretization method in the time direction in each sub-problem, which is quite useful in practice (see e.g. [17, 18, 19]).

We now explain the suitability of the SWR method. On the one hand, the analysis of this method, as well as its optimization, can be carried out directly at the (continuous) PDE level, which yields an insight which is completely independent from the actual discretization that is used in practice. On the other hand, it can be run in parallel (one processor for each sub-domain), and this allows a two-level parallelization process: one level in Parareal iterations, and the other in the SWR iterations. In addition, as we do not run the solvers until convergence, we shall need to keep additional intermediate outputs from the solvers during Parareal iterations: these outputs are fluxes on the interface. Using SWR with Dirichlet transmission condition was proposed in [16] and [5]. In this paper, we choose the Optimized transmission condition; then, the corresponding SWR method can be called Optimized Schwarz Waveform Relaxation (OSWR), see [10]. While OSWR converges much faster than Dirichlet SWR, it also amplifies the difficulties in the convergence analysis, which cannot be performed using the same techniques as for Dirichlet SWR.

In this contribution, we study as a model problem the coupled Parareal - OSWR method on one-dimensional parabolic equations. Our paper is organized as follows. In Section 2, we introduce the model problem and state some stability and regularity properties of the solutions of the whole space-time domain problem and of the sub-space-time-domain problems. In Section 3 we recall the Parareal algorithm. In Section 4 we recall and extend some known results about the OSWR method, which will be used for the analysis of the coupled method. In Section 5, we introduce the coupled Parareal-OSWR algorithm and prove a general convergence result for the method via energy estimates in Section 6. Finally, in Section 7, results of two-dimensional (2D) numerical experiments comparing the different methods (OSWR, Parareal, and the coupled Parareal-OSWR algorithm) are discussed.

## 2 Model problem

Denote  $\mathcal{L}u := \partial_t u - \nu \partial_{xx} u + a \partial_x u + bu$ ,  $\Omega = \mathbb{R}$ . For  $T > 0$ , we consider the following problem

$$\mathcal{L}(u) = f, \quad \text{in } \Omega \times (0, T), \quad (1a)$$

$$u(\cdot, 0) = u_0, \quad \text{in } \Omega, \quad (1b)$$

where  $\nu$ ,  $a$  and  $b$  are constants, with  $\nu > 0$  and  $b > 0$ . The source term  $f$  and the initial condition  $u_0$  will be specified in Section 2.2.

### 2.1 Domain decomposition and notation

We consider a decomposition of  $\Omega$  into two non-overlapping subdomains

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

and introduce the Robin interface operator for  $i = 1, 2$  as follows (see [25, 8]):

$$\mathcal{B}_1 = \nu \partial_x - \frac{a}{2} + \frac{p}{2}, \quad \mathcal{B}_2 = -\nu \partial_x + \frac{a}{2} + \frac{p}{2}, \quad (2)$$

Then, problem (1) can be reformulated as the following equivalent multi-domain problem [26], with  $f_i = f|_{\Omega_i}$ ,  $u_i = u|_{\Omega_i}$ , and  $u_{0,i} = u_0|_{\Omega_i}$ ,  $i = 1, 2$ :

$$\begin{aligned} \mathcal{L}u_i &= f_i && \text{in } \Omega_i \times (0, T), \\ u_i(\cdot, 0) &= u_{0,i} && \text{in } \Omega_i, \\ (\mathcal{B}_i u_i)(0, \cdot) &= \xi_i && \text{on } (0, T), \end{aligned} \quad i = 1, 2, \quad (3)$$

with

$$\xi_i := (\mathcal{B}_i u_j)(0, \cdot) \quad \text{on } (0, T), \quad j = 3 - i, \quad i = 1, 2. \quad (4)$$

In (2)  $p$  is a free parameter chosen such that : a) a Robin subdomain problem of type (3) is well-posed, b) it leads to a fast converging algorithm (see Section 4). Both issues will be specified later.

In what follows we will use the notation  $\boldsymbol{\xi} := (\xi_1, \xi_2)$  for the Robin data on  $(0, T)$  associated to the solution  $u$  of (1).

## 2.2 Existence and regularity results

We introduce the following spaces corresponding to the subdomain problems,

$$\mathcal{X}_i = H^1(\Omega_i), \quad i = 1, 2, \quad \mathcal{Y} = H^{\frac{1}{4}}(0, T),$$

and the local-regular space

$$\mathcal{X} := \{U \in L^2(\Omega) : U|_{\Omega_i} \in \mathcal{X}_i\},$$

equipped with the norm  $\|U\|_{\mathcal{X}} = (\sum_i \|U|_{\Omega_i}\|_{\mathcal{X}_i}^2)^{\frac{1}{2}}$ .

With the Robin transmission conditions, 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 [24].

We recall below some useful regularity properties from [25].

**Lemma 2.1.** *(Regularity of problem (1))*

If  $u_0 \in H^1(\Omega)$  and  $f \in L^2(0, T; L^2(\Omega))$ , problem (1) has a unique solution  $u$  in  $H^{2,1}(\Omega \times (0, T))$  and there exists a constant  $C$  independent of  $u_0$  and  $f$  s.t.

$$\|u\|_{H^{2,1}(\Omega \times (0, T))} \leq C(\|u_0\|_{H^1(\Omega)} + \|f\|_{L^2(0, T; L^2(\Omega))}). \quad (5)$$

**Lemma 2.2.** *(Regularity of problem (3))*

Let  $i = 1$  or  $i = 2$ . If  $u_{0,i} \in \mathcal{X}_i$ ,  $f_i \in L^2(0, T; L^2(\Omega_i))$  and  $\xi_i \in \mathcal{Y}$ , problem (3) has a unique solution  $u_i$  in  $H^{2,1}(\Omega_i \times (0, T))$  and there exists a constant  $C$  independent of  $u_0$ ,  $f$ , and  $\xi_i$  s.t.

$$\|u_i\|_{H^{2,1}(\Omega_i \times (0, T))} \leq C(\|u_{0,i}\|_{\mathcal{X}_i} + \|f_i\|_{L^2(0, T; L^2(\Omega_i))} + \|\xi_i\|_{\mathcal{Y}}).$$

**Lemma 2.3.** *(Trace theorem)*

If  $u \in H^{2,1}(\Omega \times (0, T))$ , then  $u(\cdot, T) \in \mathcal{X}$ ,  $(\mathcal{B}_i u)(0, \cdot) \in \mathcal{Y}$ ,  $i = 1, 2$ , and there exists a constant  $C$  s.t.\*

$$\|(\mathcal{B}_i u)(0, \cdot)\|_{\mathcal{Y}} \leq C\|u\|_{H^{2,1}(\Omega \times (0, T))}, \quad i = 1, 2.$$

Similar estimates hold by replacing  $\Omega$  by  $\Omega_i$ ,  $u$  by  $u_i$ , and  $\mathcal{X}$  by  $\mathcal{X}_i$ , for  $i = 1, 2$ .

## 3 Parareal Method

The Parareal method introduced in [23] is a numerical method designed to solve evolution problems in parallel. It is based on a *decomposition in time* of  $(0, T)$  into subintervals :  $(0, T) = \cup_{n=0}^{N-1} \mathcal{I}_n$ , with  $\mathcal{I}_n = (T_n, T_{n+1})$ ,  $0 \leq n \leq N - 1$ , and  $0 = T_0 < T_1 < \dots < T_{N-1} < T_N = T$ . Over each such interval generically noted as  $\mathcal{I} := (t_0, t_1)$ , it uses two propagation operators :

- $\mathcal{G}(\mathcal{I}, U_0)$  that provides a rough approximation of  $u(\cdot, t_1)$ , where  $u$  is the solution of (1), with initial condition  $u(\cdot, t_0) = U_0$ .
- $\mathcal{F}(\mathcal{I}, U_0)$  that provides a more accurate approximation of  $u(\cdot, t_1)$ .

---

\*Note that we have a better result :  $u(\cdot, T) \in H^1(\Omega)$ .

For simplicity, we will consider a regular decomposition of  $(0, T)$ , i.e. such that  $T_{n+1} - T_n = \Delta T$ . The plain Parareal algorithm [23] is as follows :

---

**Algorithm 1** (Parareal)

---

Choose an initial data  $(U_n^0)_{n \in \llbracket 0, N \rrbracket}$  with  $U_0^0 = u_0$  and  $U_n^0$  an approximation of  $u(\cdot, T_n)$ , for example :  $U_n^0 := \mathcal{G}(\mathcal{I}_{n-1}, U_{n-1}^0)$ , for  $n = 1, 2, \dots, N$ .

**for**  $k = 0, 1, \dots$  (Parareal iterations)

Set  $U_0^{k+1} = u_0$  and perform the correction iterations

$$U_{n+1}^{k+1} = \mathcal{G}(\mathcal{I}_n, U_n^{k+1}) + \mathcal{F}(\mathcal{I}_n, U_n^k) - \mathcal{G}(\mathcal{I}_n, U_n^k), \quad n = 0, 1, \dots, N-1. \quad (6)$$

**end for**

---

We denote by  $u_n$  the solution of the (sequential) fine propagator at time  $T_n$ :

$$u_n = \mathcal{F}((0, T_n), u_0), \quad u_n = \mathcal{F}(\mathcal{I}_{n-1}, u_{n-1}), \quad \forall n \in \llbracket 1, N \rrbracket.$$

In practice,  $\mathcal{F}$  will be close to exact, and thus, for the analysis presented here, we suppose that  $\mathcal{F}$  is the exact propagator, i.e.:  $\mathcal{F}((t_0, t_1), \tilde{u}_0) = \hat{u}(t_1)$  where  $\hat{u}$  is the solution of (1a) with initial condition  $\tilde{u}_0$  at  $t = t_0$ . In particular we can identify  $u_n$  with  $u(\cdot, T_n)$ , where  $u$  is the solution of (1).

## 4 Optimized Schwarz Waveform Relaxation Method

The OSWR method [9, 11, 29], for solving problem (1) is a space-time parallel method based on a *domain decomposition in space* only

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

Let  $\mathcal{I} = (0, T)$ , and  $\mathbf{I}^+$  the set of intervals of  $\mathbb{R}^+$ . The method solves iteratively subproblems on  $\Omega_1 \times \mathcal{I}$  and  $\Omega_2 \times \mathcal{I}$ , exchanging space-time boundary data<sup>†</sup> through the Robin operators  $\mathcal{B}_1$  and  $\mathcal{B}_2$  (defined in (2)), where the parameter  $p$  is chosen to optimize the convergence factor of the algorithm. The method is thus defined using a subproblem solution operator and a transmission operator: for  $i = 1, 2$  :

- the solution operator  $\mathcal{M}_i(\mathcal{I}, u_{0,i}, \xi_i)$ ,  $i = 1, 2$ , that maps the available Robin condition  $\xi_i$  and initial condition  $u_{0,i}$  to the subdomain solution  $u_i$ <sup>‡</sup>:

$$\mathcal{M}_i : \begin{array}{ccc} \mathbf{I}^+ \times H^1(\Omega_i) \times \mathcal{Y} & \rightarrow & H^{2,1}(\Omega_i \times \mathcal{I}) \\ (\mathcal{I}, u_{0,i}, \xi_i) & \rightarrow & u_i, \end{array} \quad (7)$$

where  $u_i$  is the solution of the following Robin problem in  $\Omega_i \times \mathcal{I}$

$$\begin{array}{ccc} \mathcal{L}u_i = f & \text{in } \Omega_i \times \mathcal{I}, \\ u_i(\cdot, 0) = u_{0,i} & \text{in } \Omega_i, \\ (\mathcal{B}_i u_i)(0, \cdot) = \xi_i & \text{on } \mathcal{I}. \end{array} \quad i = 1, 2, \quad (8)$$

- the transmission operator  $\mathcal{B}_i$ ,  $i = 1, 2$ , that maps the available neighbor subdomain solution  $u_j \in H^{2,1}(\Omega_j \times \mathcal{I})$ ,  $j = 3 - i$ , to a new Robin datum  $\xi_i \in \mathcal{Y}$  :  $\xi_i = (\mathcal{B}_i u_j)(0, \cdot)$  on  $\mathcal{I}$ .

Using the definition of  $\mathcal{M}_i$ , problem (3)-(4) can be rewritten as

$$u_i = \mathcal{M}_i(\mathcal{I}, u_{0,i}, \xi_i), \quad i = 1, 2, \quad (9a)$$

$$\xi_i = (\mathcal{B}_i u_j)(0, \cdot) \quad \text{on } \mathcal{I}, \quad j = 3 - i, \quad i = 1, 2. \quad (9b)$$

The OSWR algorithm for solving problem (9) (or equivalently (1)) is as follows.

---

<sup>†</sup>In the 1d case considered here, the interface is reduced to one point in space, thus the exchanged data depend on time only.

<sup>‡</sup>the operator  $\mathcal{M}_i$  should depend also on  $f$  but we omit it here to simplify the notations.

---

**Algorithm 2** (OSWR)

---

Choose an initial Robin data  $\boldsymbol{\xi}^0 = (\xi_1^0, \xi_2^0)$  on  $\mathcal{I}$ , for example  $\xi_i^0 = (\mathcal{B}_i u_{0,i})(0, \cdot)$ , on  $(0, T)$ ,  $i = 1, 2$ .  
**for**  $\ell = 1, 2, \dots$  (OSWR iterations)

1. Solve the local space-time Robin problems by calculating

$$u_i^\ell = \mathcal{M}_i(\mathcal{I}, u_{0,i}, \xi_i^{\ell-1}), \quad i = 1, 2. \quad (10)$$

2. Update the Robin interface term  $\boldsymbol{\xi}^\ell = (\xi_1^\ell, \xi_2^\ell)$ , with

$$\xi_i^\ell = (\mathcal{B}_i u_j^\ell)(0, \cdot) \quad \text{on } \mathcal{I}, \quad j = 3 - i, \quad i = 1, 2. \quad (11)$$

**end for**

---

**Remark 4.1.** By definition of  $\mathcal{M}_i$  and  $u_i^\ell$  in (10), the interface condition is

$$(\mathcal{B}_i u_i^\ell)(0, \cdot) = \xi_i^{\ell-1}, \quad i = 1, 2. \quad (12)$$

Then, from (11), and using that  $\mathcal{B}_i = -\mathcal{B}_j + p$ , we obtain, for  $\ell \geq 1$

$$\xi_i^\ell = -\xi_j^{\ell-1} + p u_j^\ell, \quad j = 3 - i, \quad i = 1, 2. \quad (13)$$

Let  $L \in \mathbb{N}^*$ . In the sequel we will denote in compact form

$$(u^L, \boldsymbol{\xi}^L) = \text{OSWR}_L(\mathcal{I}, u_0, \boldsymbol{\xi}^0), \quad (14)$$

where  $u^L \in L^2(\Omega \times \mathcal{I})$  with  $u^L|_{\Omega_i} = u_i^L$ ,  $i = 1, 2$ , and  $\boldsymbol{\xi}^L = (\xi_1^L, \xi_2^L)$  are the output after  $L$  iterations of algorithm (10)–(11) with initial condition  $u_0$  and initial Robin datum  $\boldsymbol{\xi}^0$  on  $\mathcal{I}$ .

## 4.1 Stability and convergence

We suppose that  $f \in L^2(0, T; L^2(\Omega))$ . For simplicity, we will use the notation  $\|\cdot\|$  for the  $L^2$ -norm in  $\Omega$  or in  $\Omega_i$ ,  $i = 1, 2$ , and  $\|\cdot\|_{\mathcal{I}}$  for the  $(L^2(\mathcal{I}))^2$ -norm.

Let  $(u_i^\ell, \xi_i^\ell)$ ,  $i = 1, 2$  be defined by (10)–(11). For the convergence analysis below, we introduce the following notations for the errors, for  $i = 1, 2$  and  $\ell \geq 1$ :

$$\zeta_i^\ell := \xi_i^\ell - \xi_i, \quad \text{where } \xi_i \text{ is defined in (4), and we set } \boldsymbol{\zeta}^\ell := (\zeta_1^\ell, \zeta_2^\ell), \quad (15)$$

$$e_i^\ell := u_i^\ell - u, \quad \text{where } u \text{ is the solution of (1),} \quad (16)$$

$$e^\ell := \text{the function in } L^2(\Omega \times (0, T)) \text{ s.t. } e^\ell|_{\Omega_i} = e_i^\ell, \quad i = 1, 2. \quad (17)$$

**Theorem 4.2.** Let  $L \in \mathbb{N}^*$ . If  $u_0 \in H^1(\Omega)$  and  $\boldsymbol{\xi}^0 \in \mathcal{Y}^2$ , then, Algorithm 2 is well-defined and we have

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0, T, H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\boldsymbol{\zeta}^L\|_{\mathcal{I}}^2 = \frac{1}{2p} \|\boldsymbol{\zeta}^0\|_{\mathcal{I}}^2.$$

Hence, Algorithm 2 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 (1).

*Proof.* We proceed by recurrence. We have  $u_{0,i} \in \mathcal{X}_i$  independently of  $\ell$ . Moreover, let us suppose that  $\xi_i^{\ell-1} \in \mathcal{Y}$  (this is true for  $\ell = 1$ ). Then from Lemma 2.2 we have  $u_i^\ell \in H^{2,1}(\Omega_i \times (0, T))$ . Then from Lemma 2.3, or using (13) and the Trace theorem, we have  $\xi_i^\ell \in \mathcal{Y}$ , thus Algorithm 2 is well-defined. The proof of the energy estimate is done in [8] (Theorem 5.15) and we recall it in Appendix A.  $\square$

In the context of the coupled Parareal-OSWR method in Section 5, incomplete iterations of the OSWR algorithm are performed at each Parareal iteration. This implies that the new initial condition for the OSWR algorithm, through Parareal iterations, will no more be in  $H^1(\Omega)$ , but only in  $\mathcal{X}$ . Therefore we need the following extended result, that will be used to prove the convergence of the coupled Parareal-OSWR method later.

We still use notations (15)–(17) for the errors. We suppose that the initial condition of Algorithm 2, denoted now by  $\bar{u}_0$ , verifies  $\bar{u}_0 \in \mathcal{X}$ ,  $\bar{u}_0 \neq u(\cdot, 0)$  (where  $u$  is the solution of (1)), and we introduce the additional notation for the error at time  $t = 0$  :

$$e_0 := \bar{u}_0 - u(\cdot, 0),$$

with  $e_{0,i} := e_0|_{\Omega_i}$ ,  $i = 1, 2$ .

**Theorem 4.3.** *Let  $L \in \mathbb{N}^*$ . If the initial condition of Algorithm 2 is  $\bar{u}_0 \in \mathcal{X}$ ,  $\bar{u}_0 \neq u(\cdot, 0)$ , and if  $\boldsymbol{\xi}^0 \in \mathcal{Y}^2$ , then, Algorithm 2 is well-defined and we have*

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0,T,H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\boldsymbol{\zeta}^L\|_{\mathcal{I}}^2 = \frac{L}{2} \|e_0\|^2 + \frac{1}{2p} \|\boldsymbol{\zeta}^0\|_{\mathcal{I}}^2. \quad (18)$$

*Proof.* For  $i = 1, 2$ , let  $(u_i^\ell, \xi_i^\ell)$  be defined by (10)–(11) with initial condition  $\bar{u}_0$ . We set  $\zeta_i^\ell := \xi_i^\ell - \xi_i$ , where  $\xi_i$  is defined in (4), and  $e_i^\ell := u_i^\ell - u|_{\Omega_i}$ , where  $u$  is the solution of (1).

With  $(\zeta_1^0, \zeta_2^0)$  given, the error  $e_i^\ell$ ,  $i = 1, 2$ , satisfies, for  $\ell \geq 1$  :

$$\begin{aligned} \mathcal{L}e_i^\ell &= 0 && \text{in } \Omega_i \times \mathcal{I}, \\ e_i^\ell(\cdot, 0) &= e_{0,i} && \text{in } \Omega_i, \\ (\mathcal{B}_i e_i^\ell)(0, \cdot) &= \zeta_i^{\ell-1} && \text{on } \mathcal{I}, \end{aligned} \quad i = 1, 2, \quad (19)$$

$$\text{where } \zeta_i^\ell = (\mathcal{B}_i e_j^\ell)(0, \cdot), \ell \geq 1, \quad j = 3 - i, \quad i = 1, 2. \quad (20)$$

Then we follow the same steps as in the proof of Theorem 4.2 in Appendix A until estimate (48), which is written, using that  $e_i^\ell(\cdot, 0) = e_{0,i}, \forall \ell \geq 1$  :

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0,T,H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\boldsymbol{\zeta}^L\|_{\mathcal{I}}^2 = \frac{L}{2} \sum_i \|e_{0,i}\|^2 + \frac{1}{2p} \|\boldsymbol{\zeta}^0\|_{\mathcal{I}}^2,$$

which ends the proof of Theorem 4.3. □

## 4.2 Optimized Robin parameters

In this section we give the methodology to calculate the Robin parameter  $p$  involved in the OSWR method. This parameter is chosen to optimize the convergence factor of the algorithm, and is thus called "optimized parameter".

The calculation of the convergence factor extends the one of Lemma 5.7 in [8], obtained for  $u_0 \in H^2(\Omega)$ , to the case  $u_0 \in H^1(\Omega)$ . By linearity of  $\mathcal{M}_i$ , from (9a) and (10), the error  $e_i^\ell := u_i^\ell - u$ ,  $i = 1, 2$ , at iteration  $\ell$  of the OSWR method, satisfies  $e_i^\ell = \mathcal{M}_i(\mathcal{I}, 0, \zeta_i^{\ell-1})$ , with  $\zeta_i^{\ell-1} = \xi_i^{\ell-1} - \xi_i$  and  $f = 0$ . Equivalently,  $e_i^\ell$  is solution of the following problem

$$\begin{aligned} \mathcal{L}e_i^\ell &= 0 && \text{in } \Omega_i \times \mathcal{I}, \\ e_i^\ell(\cdot, 0) &= 0 && \text{in } \Omega_i, \\ (\mathcal{B}_i e_i^\ell)(0, \cdot) &= \zeta_i^{\ell-1} && \text{on } \mathcal{I}. \end{aligned} \quad i = 1, 2, \quad (21)$$

From (9b) and (11), we have  $\zeta_i^\ell = \mathcal{B}_i e_j^\ell(0, \cdot)$ ,  $i = 1, 2$ , and thus the transmission conditions on  $\mathcal{I}$  in (21) also reads

$$(\mathcal{B}_i e_i^\ell)(0, \cdot) = (\mathcal{B}_i e_j^{\ell-1})(0, \cdot) \quad \text{on } \mathcal{I}, \quad j = 3 - i, \quad i = 1, 2. \quad (22)$$

Now, in order to use the Fourier transform, we extend (21)–(22) to  $\mathbb{R}$  in the following way :

- for  $\ell = 1$ , the Robin condition in (21) is with  $\zeta_i^0 \in \mathcal{Y}$  and we can extend  $\zeta_i^0$  by zero to  $H^{\frac{1}{4}}(\mathbb{R})$  to obtain a function, denoted by  $\tilde{\zeta}_i^0$ , vanishing on  $(-\infty, 0)$ , and on  $(T, +\infty)$ , for  $i = 1, 2$ . Then we can extend equations (21) to  $\mathbb{R}$ , and their solutions, denoted by  $(\tilde{e}_1^1, \tilde{e}_2^1)$ , vanish on  $(-\infty, 0)$  and coincide with  $(e_1^1, e_2^1)$  on  $(0, T)$ .

• for  $\ell \geq 2$ , we define the Robin trace  $\tilde{\zeta}_i^{\ell-1} := (\mathcal{B}_i \tilde{e}_j^{\ell-1})(0, \cdot)$  on  $\{0\} \times \mathbb{R}$ , which belongs to  $H^{\frac{1}{4}}(\mathbb{R})$ , vanishes on  $(-\infty, 0)$  and coincides with  $(\mathcal{B}_i e_j^{\ell-1})(0, \cdot)$  on  $\{0\} \times (0, T)$ , for  $i = 1, 2$ . The subdomain problems (21) are extended on  $\Omega_i \times \mathbb{R}$  as follows :

$$\begin{aligned} \mathcal{L} \tilde{e}_i^\ell &= 0 && \text{in } \Omega_i \times \mathbb{R}, \\ \tilde{e}_i^\ell(\cdot, 0) &= 0 && \text{in } \Omega_i, \\ (\mathcal{B}_i \tilde{e}_i^\ell)(0, \cdot) &= \tilde{\zeta}_i^{\ell-1} && \text{on } \mathbb{R}, \end{aligned} \quad i = 1, 2, \quad (23)$$

and their solution vanish on  $(-\infty, 0)$  and coincide with  $(e_1^\ell, e_2^\ell)$  on  $(0, T)$ . In particular, by the definition of  $\tilde{\zeta}_i^{\ell-1}$  above, the transmission condition (22) has been extended on  $\{0\} \times \mathbb{R}$  in (23) as follows :

$$(\mathcal{B}_i \tilde{e}_i^\ell)(0, \cdot) = (\mathcal{B}_i \tilde{e}_j^{\ell-1})(0, \cdot) \quad \text{on } \mathbb{R}, \quad j = 3 - i, \quad i = 1, 2. \quad (24)$$

Note that for  $\ell \geq 1$ , the property  $\tilde{\zeta}_i^{\ell-1} \in H^{\frac{1}{4}}(\mathbb{R})$  implies that the solution  $e_i^\ell$  of problem (23) is in  $H_{\text{loc}}^{2,1}(\Omega_i \times \mathbb{R})$ <sup>§</sup> (using Lemma 2.2), and thus the new Robin datum  $\tilde{\zeta}_i^\ell$ , defined above, is in  $H_{\text{loc}}^{\frac{1}{4}}(\mathbb{R})$  (using Lemma 2.3).

In what follows, we use Fourier transform in time, in the sense of tempered distributions. Then, we solve in each subdomain the ordinary differential equation

$$i\omega \hat{e}_i^\ell - \nu \partial_{xx} \hat{e}_i^\ell + a \partial_x \hat{e}_i^\ell + b \hat{e}_i^\ell = 0, \quad i = 1, 2, \quad (25)$$

with the characteristic roots

$$r^- = \frac{a - \sqrt{d}}{2\nu}, \quad r^+ = \frac{a + \sqrt{d}}{2\nu}, \quad d = a^2 + 4\nu(b + i\omega), \quad (26)$$

where  $\sqrt{d}$  is the complex square-root with positive real part : let  $\tilde{d} = \sqrt{(a^2 + 4\nu b)^2 + 16\nu^2 \omega^2}$ , then

$$\sqrt{d} = \sqrt{\frac{\tilde{d} + a^2 + 4\nu b}{2}} + i \operatorname{sign}(\omega) \sqrt{\frac{\tilde{d} - a^2 - 4\nu b}{2}}.$$

Thus,  $\mathcal{R}e(r^+) > 0$  and  $\mathcal{R}e(r^-) < 0$ , and the solutions  $\hat{e}_i^\ell \in L^2(\Omega_i)$ ,  $i = 1, 2$  are given by<sup>¶</sup>

$$\hat{e}_1^\ell = \frac{2}{\sqrt{d} + p} \hat{\zeta}_1^{\ell-1}(\omega) e^{r^+ x}, \quad \hat{e}_2^\ell = \frac{2}{\sqrt{d} + p} \hat{\zeta}_2^{\ell-1}(\omega) e^{r^- x}, \quad \ell \geq 1. \quad (27)$$

Then, replacing (27) in the transmission conditions (24) leads to

$$\forall \ell \geq 1, \quad \begin{pmatrix} \hat{\zeta}_1^\ell \\ \hat{\zeta}_2^\ell \end{pmatrix} = \begin{pmatrix} -\sqrt{d} + p \\ \sqrt{d} + p \end{pmatrix} \begin{pmatrix} \hat{\zeta}_2^{\ell-1} \\ \hat{\zeta}_1^{\ell-1} \end{pmatrix}. \quad (28)$$

Setting  $\hat{\zeta}^\ell := (\hat{\zeta}_1^\ell, \hat{\zeta}_2^\ell)$ , from (28), we have

$$\forall \ell \geq 2, \quad \hat{\zeta}^\ell = \rho_0(\omega, p) \hat{\zeta}^{\ell-2}, \quad \text{with } \rho_0(\omega, p) := \left( \frac{-\sqrt{d} + p}{\sqrt{d} + p} \right)^2. \quad (29)$$

From (29), by induction on  $\ell$  we obtain,

$$\hat{\zeta}^{2\ell} = (\rho_0(\omega, p))^\ell \hat{\zeta}^{(0)}, \quad \forall \ell \geq 1. \quad (30)$$

<sup>§</sup>Here we set  $H_{\text{loc}}^{2,1}(\Omega_i \times \mathbb{R}) = \{H^{2,1}(\Omega_i \times (T_1, T_2)), \forall T_1, T_2 \in \mathbb{R}\}$ ,  $i = 1, 2$ .

<sup>¶</sup>Note that here the term “ $\partial_x$ ” in  $\mathcal{B}_i$  is multiplied by  $\nu$  while this is not the case in [8]. Thus  $\hat{e}_i^\ell$  is slightly different here, from the one of [8].



From (30), the convergence factor of the algorithm is  $\rho_0(\omega, p)$ , defined in (29).

Note that from (28) we have, for all  $\omega \in \mathbb{R}$ , and for  $\ell \geq 1$

$$|\hat{\zeta}_i^\ell(\omega)| \leq |\hat{\zeta}_{3-i}^{\ell-1}(\omega)|, \quad \text{for } i = 1, 2.$$

Using that  $\tilde{\zeta}^0 \in H^{\frac{1}{4}}(\mathbb{R})$ , the above inequality implies, by induction, that

$$\tilde{\zeta}^\ell \in H^{\frac{1}{4}}(\mathbb{R}), \quad \forall \ell \geq 1.$$

While we have  $\max_{\omega \in \mathbb{R}} |\rho_0(\omega, p)| = \lim_{\omega \rightarrow \infty} |\rho_0(\omega, p)| = 1$ , we can use the continuous convergence factor  $\rho_0(\omega, p)$  to calculate an efficient Robin parameter for the discrete setting (see e.g. [21, 29, 6, 8]).

Indeed, in numerical computations, the frequency  $\omega$  is bounded, i.e. we have  $\omega_{\min} \leq \omega \leq \omega_{\max}$  where  $\omega_{\max} = \frac{\pi}{\Delta t}$  is the largest discrete frequency supported by the numerical time grid, and  $\omega_{\min} = \frac{\pi}{T}$  is smallest frequency relevant to the global time interval. Defining

$$\rho_c(p) := \max_{\frac{\pi}{T} \leq \omega \leq \frac{\pi}{\Delta t}} |\rho_0(p, \omega)|,$$

then the optimized Robin parameter  $p_c$  is chosen such that it verifies

$$\rho_c(p_c) = \min_{p > 0} \rho_c(p). \quad (31)$$

In practice, the minimization problem (31) is solved numerically, using the `fminsearch` function in MATLAB.

## 5 Coupled Parareal-OSWR method

Coupling Parareal with domain decomposition was introduced in [16, 12, 5].

Like in Section 3, we set  $(0, T) = \cup_{n=0}^{N-1} \mathcal{I}_n$ . Then the Parareal-OSWR algorithm is defined using the coarse propagator  $\mathcal{G}$  of Section 3 and the incomplete<sup>||</sup> fine propagator  $\text{OSWR}_L$  defined by (14) as follows :

---

### Algorithm 3 (Coupled Parareal-OSWR)

---

1. Choose an initial data  $(U_n^0)_{n \in [0, N]}$  with  $U_0^0 = u_0$  and  $U_n^0$  an approximation of  $u(\cdot, T_n)$ , for example  $U_n^0 := \mathcal{G}(\mathcal{I}_{n-1}, U_{n-1}^0)$ , for  $n = 1, 2, \dots, N$ .
2. Choose an initial Robin data  $(\boldsymbol{\xi}_n^{0,0})_{n \in [0, N-1]}$ , with  $\boldsymbol{\xi}_n^{0,0} := (\xi_{1,n}^{0,0}, \xi_{2,n}^{0,0})$  on  $\mathcal{I}_n$ , for example  $\xi_{i,n}^{0,0} = (\mathcal{B}_i U_n^0)(0, \cdot)$ ,  $i = 1, 2$ .

**for**  $k = 0, 1, \dots$  (*Parareal iterations*)

1. On each time interval  $\mathcal{I}_n$ ,  $n = 0, 1, \dots, N-1$  :

$$\text{Calculate } (u_n^{k,L}, \boldsymbol{\xi}_n^{k,L}) = \text{OSWR}_L(\mathcal{I}_n, U_n^k, \boldsymbol{\xi}_n^{k,0}). \quad (32)$$

2. Set  $U_0^{k+1} = u_0$  and do Parareal corrections:

$$U_{n+1}^{k+1} = u_n^{k,L}(\cdot, T_{n+1}) + \mathcal{G}(\mathcal{I}_n, U_n^{k+1}) - \mathcal{G}(\mathcal{I}_n, U_n^k). \quad (33)$$

$$\text{Update the interface term: } \boldsymbol{\xi}_n^{k+1,0} = \boldsymbol{\xi}_n^{k,L}. \quad (34)$$

**end for**

---

**Remark 5.1.** In (32), if  $L = \infty$ , then  $u_n^{k,L} \in H^1(\Omega)$ , for all  $n \geq 0$ ,  $k \geq 0$ . However, if  $L < \infty$  and chosen small (e.g.  $L = 2$ ), then  $u_n^{k,L} \in \mathcal{X}$ , for all  $n \geq 0$ ,  $k \geq 0$ , but there is very few chance that  $u_n^{k,L}$  be in  $H^1(\Omega)$ . Thus, using (33) and that  $u_0 \in H^1(\Omega) \subset \mathcal{X}$ , we will have  $U_n^k \in \mathcal{X}$ , for all  $n \geq 0$ ,  $k \geq 0$ .

<sup>||</sup>In the sense that  $L$  will be smaller than the number of iterations required for convergence.

## 6 Convergence of the Parareal-OSWR algorithm

We will consider the convergence in the  $L^2(\Omega)$ -norm. As in Section 4,  $\|\cdot\|$  will stand for the  $L^2$ -norm in  $\Omega$  or in  $\Omega_i$ ,  $i = 1, 2$ , and  $\|\cdot\|_{\mathcal{I}}$  the  $(L^2(\mathcal{I}))^2$ -norm. Let  $\tilde{\mathcal{G}}$  be the coarse propagator associated to the source term  $f = 0$ . We have the following result

**Theorem 6.1.** *We suppose that there exists a constant  $\gamma_1$  such that  $\tilde{\mathcal{G}}$  satisfies for each  $n = 0, 1, \dots, N$*

$$\|\tilde{\mathcal{G}}(\mathcal{I}_n, U)\| \leq \gamma_1 \|U\|, \quad \forall U \in L^2(\Omega). \quad (35)$$

Then, when  $k \rightarrow \infty$

- $U_n^k$  converges to  $u(\cdot, T_n)$  in  $L^2(\Omega)$ ,
- $u_{i,n}^{k,\ell}$  converges to  $u|_{\Omega_i \times [T_n, T_{n+1}]}$  in  $L^2(T_n, T_{n+1}, H^1(\Omega))$  for all  $\ell = 1, 2, \dots, L$ .

In order to prove Theorem 6.1, we first introduce some notation and prove the following lemmas.

### Notation for error estimation

Let  $u$  be the solution of problem (1), and  $(u_n^{k,\ell}, \boldsymbol{\xi}_n^{k,\ell})_{1 \leq \ell \leq L}$  be the sequence of iterates through the OSWR step (32). We define, for  $k = 0, 1, \dots$ , and  $n = 0, \dots, N-1$ :

- $E_n^k := U_n^k - u_n$ ,  $n = 0, \dots, N$ , where  $u_n = u(\cdot, T_n)$ ,
- $e_n^{k,\ell} := u_n^{k,\ell} - u$ , the error in  $L^2(\Omega \times \mathcal{I}_n)$  at each iteration  $\ell$  inside step (32), with  $e_{i,n}^{k,\ell} := e_n^{k,\ell}|_{\Omega_i}$ ,  $i = 1, 2$ , for  $\ell = 1, \dots, L$ ,
- $\boldsymbol{\zeta}_n^{k,\ell} := \boldsymbol{\xi}_n^{k,\ell} - \boldsymbol{\xi}_n$ , for  $\ell = 1, \dots, L$ , where  $\boldsymbol{\xi}_n = ((\mathcal{B}_1 u)(0, \mathcal{I}_n), (\mathcal{B}_2 u)(0, \mathcal{I}_n))$ .

Then, by linearity, the algorithm on the error reads

---

#### Algorithm 4 (Coupled Parareal-OSWR algorithm on the error)

---

1. Define initial data  $(E_n^0)_{n \in \llbracket 0, N \rrbracket}$  with  $E_0^0 = 0$ ,  $E_n^0 := \mathcal{G}(\mathcal{I}_{n-1}, U_{n-1}^0) - u_n$ , where  $U_0^0 = u_0$ ,  $n = 1, 2, \dots, N$ .
2. Define initial Robin data  $(\boldsymbol{\zeta}_n^{0,0})_{n \in \llbracket 0, N \rrbracket}$ , with  $\boldsymbol{\zeta}_n^{0,0} := (\boldsymbol{\zeta}_{1,n}^{0,0}, \boldsymbol{\zeta}_{2,n}^{0,0})$  on  $\mathcal{I}_n$ , where  $\boldsymbol{\zeta}_{i,n}^{0,0} = (\mathcal{B}_i E_n^0)(0, \cdot)$ ,  $i = 1, 2$ .

for  $k = 0, 1, \dots$  (Parareal iterations)

1. On each time interval  $\mathcal{I}_n$ ,  $n = 0, 1, \dots, N-1$ :

$$\text{Calculate } (e_n^{k,L}, \boldsymbol{\zeta}_n^{k,L}) = \text{OSWR}_L(\mathcal{I}_n, E_n^k, \boldsymbol{\zeta}_n^{k,0}). \quad (36)$$

2. Set  $E_n^{k+1} = 0$  and do Parareal correction:

$$E_{n+1}^{k+1} = e_n^{k,L}(\cdot, T_{n+1}) + \tilde{\mathcal{G}}(\mathcal{I}_n, E_n^{k+1} - E_n^k). \quad (37)$$

$$\text{Update the interface term: } \boldsymbol{\zeta}_n^{k+1,0} = \boldsymbol{\zeta}_n^{k,L}. \quad (38)$$

end for

---

**Lemma 6.2.** *For all  $n \in \llbracket 0, N-1 \rrbracket$ , we have:*

$$\sum_{k=0}^K \|e_n^{k,L}(\cdot, T_{n+1})\|^2 \leq L \sum_{k=0}^K \|E_n^k\|^2 + \frac{1}{p} \|\boldsymbol{\zeta}_n^{0,0}\|_{\mathcal{I}_n}^2. \quad (39)$$

*Proof.* Step (36) corresponds to the OSWR algorithm, on  $\mathcal{I} := \mathcal{I}_n$ , with initial condition  $E_n^k \in \mathcal{X}$  (see Remark 5.1). Thus, from (18) with  $e_0 := E_n^k$ ,  $e_i^\ell := e_{i,n}^{k,\ell}$ , and  $\boldsymbol{\zeta}^\ell := \boldsymbol{\zeta}_n^{k,\ell}$ , we obtain

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e_n^{k,\ell}(\cdot, T_{n+1})\|^2 + \sum_i \|e_{i,n}^{k,\ell}\|_{L^2(T_n, T_{n+1}; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\boldsymbol{\zeta}_n^{k,L}\|_{\mathcal{I}_n}^2 = \frac{L}{2} \|E_n^k\|^2 + \frac{1}{2p} \|\boldsymbol{\zeta}_n^{k,0}\|_{\mathcal{I}_n}^2.$$

Moreover, in our algorithm, from (38) we have  $\zeta_n^{k,L} = \zeta_n^{k+1,0}$ , thus

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e_n^{k,\ell}(\cdot, T_{n+1})\|^2 + \sum_i \|e_{i,n}^{k,\ell}\|_{L^2(T_n, T_{n+1}; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\zeta_n^{k+1,0}\|_{\mathcal{I}_n}^2 = \frac{L}{2} \|E_n^k\|^2 + \frac{1}{2p} \|\zeta_n^{k,0}\|_{\mathcal{I}_n}^2.$$

Summing with respect to  $k$ , from 0 to  $K$ , we get a telescopic sum on the interface and therefore

$$\begin{aligned} \sum_{k=0}^K \sum_{\ell=1}^L \left( \frac{1}{2} \|e_n^{k,\ell}(\cdot, T_{n+1})\|^2 + \sum_i \|e_{i,n}^{k,\ell}\|_{L^2(T_n, T_{n+1}; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\zeta_n^{K+1,0}\|_{\mathcal{I}_n}^2 \\ = \frac{L}{2} \sum_{k=0}^K \|E_n^k\|^2 + \frac{1}{2p} \|\zeta_n^{0,0}\|_{\mathcal{I}_n}^2, \end{aligned} \quad (40)$$

from which we obtain (39).  $\square$

**Lemma 6.3.** *We suppose that  $\tilde{\mathcal{G}}$  satisfies (35). Then, for all  $n \in \llbracket 0, N-1 \rrbracket$ ,*

$$\sum_{k=0}^K \|E_{n+1}^{k+1}\|^2 \leq 8\gamma_1^2 \sum_{k=0}^{K+1} \|E_n^k\|^2 + 2 \sum_{k=0}^K \|e_n^{k,L}(\cdot, T_{n+1})\|^2. \quad (41)$$

*Proof.* Using the triangle inequality in (37), and then (35), we get

$$\begin{aligned} \|E_{n+1}^{k+1}\|^2 &\leq 2\|\tilde{\mathcal{G}}(\mathcal{I}_n, E_n^{k+1} - E_n^k)\|^2 + 2\|e_n^{k,L}(\cdot, T_{n+1})\|^2, \\ &\leq 2\gamma_1^2 \|E_n^{k+1} - E_n^k\|^2 + 2\|e_n^{k,L}(\cdot, T_{n+1})\|^2, \\ &\leq 4\gamma_1^2 (\|E_n^{k+1}\|^2 + \|E_n^k\|^2) + 2\|e_n^{k,L}(\cdot, T_{n+1})\|^2. \end{aligned}$$

Then, summing with respect to  $k$ , from 0 to  $K$ , we have

$$\sum_{k=0}^K \|E_{n+1}^{k+1}\|^2 \leq 4\gamma_1^2 \sum_{k=0}^K (\|E_n^{k+1}\|^2 + \|E_n^k\|^2) + 2 \sum_{k=0}^K \|e_n^{k,L}(\cdot, T_{n+1})\|^2,$$

from which we deduce (41).  $\square$

With these lemmas, we can now prove Theorem 6.1

*Proof of Theorem 6.1.* From Lemmas 6.2 and 6.3, we get

$$\sum_{k=0}^K \|E_{n+1}^{k+1}\|^2 \leq 8\gamma_1^2 \sum_{k=0}^{K+1} \|E_n^k\|^2 + 2L \sum_{k=0}^K \|E_n^k\|^2 + \frac{2}{p} \|\zeta_n^{0,0}\|_{\mathcal{I}_n}^2.$$

Denoting  $\gamma_2 = 8\gamma_1^2 + 2L$  and  $R_n = \|E_{n+1}^0\|^2 + \frac{2}{p} \|\zeta_n^{0,0}\|_{\mathcal{I}_n}^2$ , we can rewrite this inequality as

$$\sum_{k=0}^{K+1} \|E_{n+1}^k\|^2 \leq \gamma_2 \sum_{k=0}^{K+1} \|E_n^k\|^2 + R_n, \quad n \in \llbracket 0, N-1 \rrbracket.$$

From this inequality, by induction, we obtain

$$\sum_{k=0}^{K+1} \|E_{n+1}^k\|^2 \leq \gamma_2^{n+1} \sum_{k=0}^{K+1} \|E_0^k\|^2 + \sum_{j=0}^n \gamma_2^j R_{n-j}, \quad n \in \llbracket 0, N-1 \rrbracket.$$

Using that  $E_0^k = 0, \forall k \geq 0$ , we finally obtain

$$\sum_{k=0}^{K+1} \|E_{n+1}^k\|^2 \leq \sum_{j=0}^n \gamma_2^j R_{n-j}, \quad n \in \llbracket 0, N-1 \rrbracket. \quad (42)$$

Since the right-hand side of (42) does not depend on  $K$ , this shows that, for a given  $n \in \llbracket 0, N \rrbracket$ , the sum  $\sum_{k=0}^K \|E_n^k\|^2$  is bounded with respect to  $K$ . Hence  $E_n^k$  converges to 0 in the  $L^2(\Omega)$ -norm when  $k \rightarrow \infty$ . Moreover, from inequality (40), we obtain, for all  $\ell = 1, 2, \dots, L$

$$\sum_{k=0}^K \|e_{i,n}^{k,\ell}\|_{L^2(T_n, T_{n+1}; H^1(\Omega_i))} \leq \frac{L}{2} \sum_{k=0}^K \|E_n^k\|^2 + \frac{1}{2p} \|\zeta_n^{0,0}\|_{\mathcal{I}_n}^2.$$

As the sum in the right-hand side of the above inequality is bounded with respect to  $K$ , then the sum in the left is also bounded with respect to  $K$ . Hence,  $e_{i,n}^{k,\ell}$  tends to 0 in  $L^2(T_n, T_{n+1}; H^1(\Omega_i))$ , i.e.  $u_{i,n}^{k,\ell} \rightarrow u_{i,n}$  in  $L^2(T_n, T_{n+1}; H^1(\Omega_i))$ , for any  $\ell$ .  $\square$

**Remark 6.4.** *The proof of convergence of the nonoverlapping Parareal-OSWR algorithm in Theorem 6.1 (using energy estimates) is done with  $U_n^k \in \mathcal{X}$ , for all  $n \geq 0, k \geq 0$  (see Remark (5.1)). This result is obtained without any correction on  $U_n^k$  so that it is more regular (i.e. in  $H^1(\Omega)$ ). Thus, in practice we do not need a correction on  $U_n^k$  to have a convergent algorithm.*

## 7 Numerical results in 2D

In this section, we give some numerical illustrations of the performances of the coupled Parareal-OSWR method (Algorithm 3), in two space dimensions. In Section 7.1 and 7.2, these results are shown in the context of the NICEM method [20] for the space discretization, that enables the use of more efficient transmission conditions in the OSWR method (i.e. Ventcel conditions) with general domain decomposition and meshes. In section 7.3 we consider mixed finite elements as in [2] and Robin transmission condition. In all cases the backward Euler scheme in time is used.

In algorithm 3, a possible choice for the the initial Robin datum on  $\mathcal{I}_n$  is to take  $\xi_{i,n}^{0,0} = (\mathcal{B}_i U_n^0)(0)$ ,  $i = 1, 2$  (i.e.  $\xi_{i,n}^{0,0}$  is constant on  $\mathcal{I}_n$ ). A better choice, that improves the convergence of the Parareal-OSWR method, is to define  $V_n^0$  as a linear interpolation between  $U_n^0$  and  $U_{n+1}^0$ , and then to take  $\xi_{i,n}^{0,0} = (\mathcal{B}_i V_n^0)(0)$ ,  $i = 1, 2$ . Thus the latter case will be considered here.

The multi-domain problem (9) can actually be reformulated as an interface problem (see [4] or [18]) that can be solved by various iterative methods, such as block-Jacobi (which corresponds to the OSWR algorithm) or GMRES, the latter being our choice here (we call it ‘‘OSWRG’’ in what follows). Thus  $L$  will designate the number of GMRES iterations. In the Parareal algorithm, both the coarse and the fine solvers are performed using the OSWRG algorithm. For the coarse solver, as well as for the fine solver with  $L = \infty$ , the stopping criterion is when the jump of the optimized transmission conditions, measured in the  $L^2$ -norm on the interface, has been reduced below  $10^{-13}$ . Otherwise, we will consider  $L$  iterations for the fine solver, with different values of  $L$ . Note that in what follows the error is measured in the  $L^\infty(0, T; H^1(\Omega))$ -norm, and that we obtain similar results in the  $L^\infty(\Omega \times (0, T))$ -norm.

### 7.1 A rotating velocity

We set  $\Omega = ]0, 1[ \times ]0, 1[$ ,  $T = 21$ , and consider the two dimensional problem

$$\partial_t u + \nabla \cdot (\mathbf{a}u) - \nu \Delta u = f, \quad \text{in } \Omega \times (0, T), \quad (43a)$$

$$u = u_D, \quad \text{on } \partial\Omega \times (0, T), \quad (43b)$$

$$u(\cdot, 0) = u_0, \quad \text{in } \Omega, \quad (43c)$$

with a rotating velocity field  $\mathbf{a} = (a_x, a_y)$ , where

$$a_x = -\sin\left(\pi\left(y - \frac{1}{2}\right)\right) \cos\left(\pi\left(x - \frac{1}{2}\right)\right), \quad a_y = \cos\left(\pi\left(y - \frac{1}{2}\right)\right) \sin\left(\pi\left(x - \frac{1}{2}\right)\right),$$

see Figure 1 on the left. We choose the right-hand side  $f$  and the values of the boundary and initial conditions so that the exact solution is given by

$$u(x, y) = \cos(\pi x) \sin(\pi y) \cos\left(\frac{2\pi t}{11}\right), \quad \forall (x, y) \in \Omega, \quad \forall t \in (0, T). \quad (44)$$

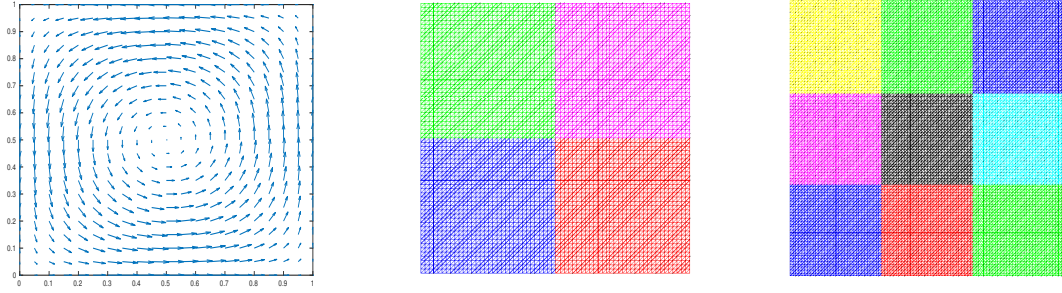


Figure 1: Example 1: rotating velocity (left), decomposition of  $\Omega$  into 4 subdomains (middle), and decomposition of  $\Omega$  into 9 subdomains (right)

The number of time windows for the Parareal iterations is  $N = 21$ . The time steps of the coarse and fine solvers are  $\Delta t_C = 1$  and  $\Delta t_F = 0.0156$ , respectively.

In what follows we denote by  $e$  the relative scheme error, in  $L^\infty(0, T; H^1(\Omega))$ -norm, between the converged Parareal solution and the solution  $u$  of problem (43), given in (44). The term ‘‘OSWRG iterations’’ will designate the iterations for the fine solver. We will consider two cases for the diffusion and the domain decomposition in space.

In the results below, the cost of the coarse solver is negligible.

### Case 1

We take  $\nu = 0.05$  and a decomposition of  $\Omega$  into four subdomains as in Figure 1 in the middle. The number of triangles in the whole domain  $\bar{\Omega}$  is 8192 and the mesh size in each subdomain is 0.0221.

In Figure 2 (left), we plot the evolution of the relative error in the  $L^\infty(0, T; H^1(\Omega))$ -norm, between the Parareal-OSWRG solution and the converged Parareal solution, as a function of the number of Parareal iterations, for different values of  $L$ . The case  $L = \infty$  (black circle curve) corresponds to  $L = 10$  (average value). The horizontal dashed green line represents  $0.1e$  and we consider the number of iterations such that the algorithm error is smaller than this value. This is obtained after 3 iterations for Parareal or Parareal-OSWRG with  $L = 8$ , and after 5, 8, 18 iterations for Parareal-OSWRG with  $L = 4$ ,  $L = 2$ ,  $L = 1$ , respectively. These data are reported in Table 1. We observe that the fastest case is  $L = 2$  with 16 OSWRG iterations globally.

Table 2 shows the gain (in term of fine solver iterations), when the fine solvers are performed in parallel, of the full Parareal ( $L = 10$ ) or the coupled Parareal-OSWRG methods compared to the OSWRG algorithm for solving problem (43). The first column corresponds to the number of OSWRG iterations to reach  $0.1e$ , when the OSWRG method is used to solve (43). We observe that for  $N = 21$  time windows, we gain a factor 10.5 for Parareal and a factor 20 for Parareal-OSWRG which is nearly the expected parallel efficiency.

L	1	2	4	8	10 ( $\approx \infty$ )
$k$	17	8	5	3	3
$L * k$	17	16	20	24	30

Table 1: Example 1 (case 1): Number of Parareal iterations  $k$  and total number of OSWRG iterations  $L * k$ , versus  $L$

### Case 2

We take  $\nu = 0.1$  and a decomposition of  $\Omega$  into nine subdomains as in Figure 1 on the right. The number of triangles in the whole domain  $\bar{\Omega}$  is 18432 and the mesh size in each subdomain is 0.0147.

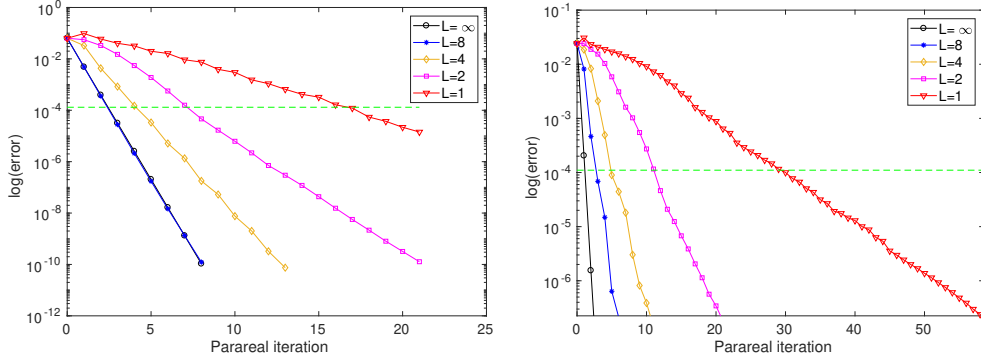


Figure 2: Example 1: Relative error versus Parareal iterations. Case  $\nu = 0.05$  with 4 subdomains (left) and case  $\nu = 0.1$  with 9 subdomains (right)

Solver	OSWRG	Parareal	Parareal-OSWRG ( $L = 2$ )
iterations ( $n$ )	15	30	16
loss factor ( $\ell = n/15$ )	–	2	1.07
gain factor ( $N$ )	–	21	21
final gain factor ( $N/\ell$ )	–	10.5	19.69

Table 2: Example 1 (case 1): Gain factor for Parareal and coupled Parareal-OSWRG methods compared to the OSWRG solver for solving problem (43)

In Figure 2 (right), the Parareal algorithm (black circle curve) corresponds to  $L = 24$  (average value). The horizontal dashed green line represents  $0.1e$ , and is reached after 2 iterations for Parareal, and after 3, 5, 12, 30 iterations for Parareal-OSWRG with  $L = 8$ ,  $L = 4$ ,  $L = 2$ ,  $L = 1$ , respectively. These data are shown in Table 3 and we observe that the fastest case is  $L = 4$ , with a total of 20 OSWRG iterations.

Table 4 shows the gain, when the fine solvers are performed in parallel, of the Parareal or the coupled Parareal-OSWRG methods compared to the OSWRG algorithm. In the first column we give the number of OSWRG iterations to reach  $0.1e$ , when it is used as a solver for problem (43). We observe that for  $N = 21$  time windows, we gain a factor 10.94 for Parareal and a factor 26.25 for Parareal-OSWRG which represents an efficiency strictly greater than 1, since the number of processors is 21. This is notably better than the expected parallel efficiency.

L	1	2	4	8	24 ( $\approx \infty$ )
$k$	30	12	5	3	2
$L * k$	30	24	20	24	48

Table 3: Example 1 (case 2): Number of Parareal iterations  $k$  and total number of OSWRG iterations  $L * k$ , versus  $L$

## 7.2 A boundary layer case with vortices

We consider problem (43) with  $f = 0$ ,  $u_0(x, y) = 1 - x$ ,  $u_D(x, y) = 1$  on  $\{x = 0\}$  and  $u_D(x, y) = 0$  elsewhere, and the following velocity field proposed in [31]:  $\mathbf{a} = 0.32\pi (\sin(4\pi x) \sin(4\pi y), \cos(4\pi x) \cos(4\pi y))$ , see Figure 3 on the left. The diffusion coefficient is  $\nu = 0.01$  and the final time is  $T = 51$ . The number of time windows for the Parareal iterations is  $N = 51$ . The time step of the coarse solver is  $\Delta t_C = 1$ .

In what follows we consider a decomposition of  $\Omega$  into nine subdomains and consider a uniform mesh

Solver	OSWRG	Parareal	Parareal-OSWRG ( $L = 4$ )
iterations ( $n$ )	25	48	20
loss factor ( $\ell = n/25$ )	–	1.92	0.8
gain factor ( $N$ )	–	21	21
final gain factor ( $N/\ell$ )	–	10.94	26.25

Table 4: Example 1 (case 2): Gain factor for Parareal and coupled Parareal-OSWRG methods compared to the OSWRG solver for solving problem (43)

(case 1) and then a nonconforming mesh adapted to the physics (case 2). The computed Parareal-OSWRG solution at final time  $t = T$  (with the nonconforming mesh) is shown in Figure 3 on the right.

### Case 1

We consider a uniform mesh as in Figure 1 on the right, with a mesh size in each subdomain equal to 0.0147. The time step of the fine solver is  $\Delta t_F = 0.0156$ .

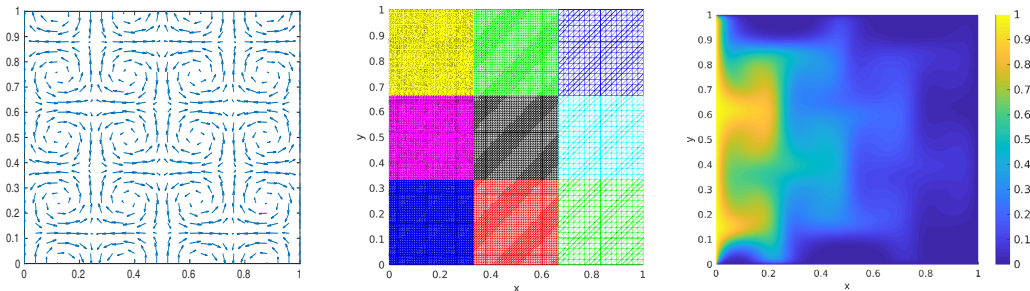


Figure 3: Example 2: velocity field (left), nonconforming meshes (middle), and computed Parareal-OSWRG solution at final time (right)

In Figure 4, we plot the evolution of the relative error in  $L^\infty(0, T; H^1(\Omega))$ -norm, between the Parareal-OSWRG solution and the converged Parareal solution, as a function of the number of Parareal iterations, for different values of  $L$ . The Parareal algorithm corresponds to  $L = 10$  and we will compare the other cases of  $L$  to this case. We do not know the solution of problem (43) and thus the relative scheme error  $e$ , however we expect that  $e$  is between  $10^{-1}$  and  $10^{-3}$ , and thus  $0.1e$  between  $10^{-2}$  and  $10^{-4}$ . Consequently, in Figure 4, the horizontal dashed green lines represent three possible values for  $0.1e$ , and Table 5 shows the number of Parareal and total OSWRG iterations to reach these different values, for  $L = 1, 2, 4, 8, 10$ . We observe that the fastest case is  $L = 2$ .

The OSWRG algorithm used as a solver for problem (43) will need 22, 36 and 49 iterations to get a relative error in  $L^\infty(0, T; H^1(\Omega))$ -norm (between the iterate solution and the converged OSWRG solution) smaller than  $10^{-2}$ ,  $10^{-3}$  and  $10^{-4}$ , respectively. Table 6 shows the gain, when the fine solvers are performed in parallel, of the Parareal or the coupled Parareal-OSWRG methods compared to the OSWRG algorithm. In this table we take the values obtained for  $e = 10^{-2}$  (note that these values correspond also to those obtained with a mean value of the values for  $e = 10^{-1}$ ,  $e = 10^{-2}$  and  $e = 10^{-3}$ ). We observe that for  $N = 51$  time windows, we gain approximately a factor 26.29 for Parareal and a factor 57.38 for Parareal-OSWRG which is slightly better than the expected parallel efficiency.

### Case 2

We consider nonconforming meshes, refined in the region of the boundary layer, see Figure 4 on the middle, with a mesh size equal to 0.0065, 0.0131, and 0.0295, for the subdomains with a boundary along  $\{x = 0\}$ , those with boundaries along  $\{x = \frac{1}{3}\}$  and  $\{x = \frac{2}{3}\}$ , and those with a boundary along  $\{x = 1\}$ , respectively.

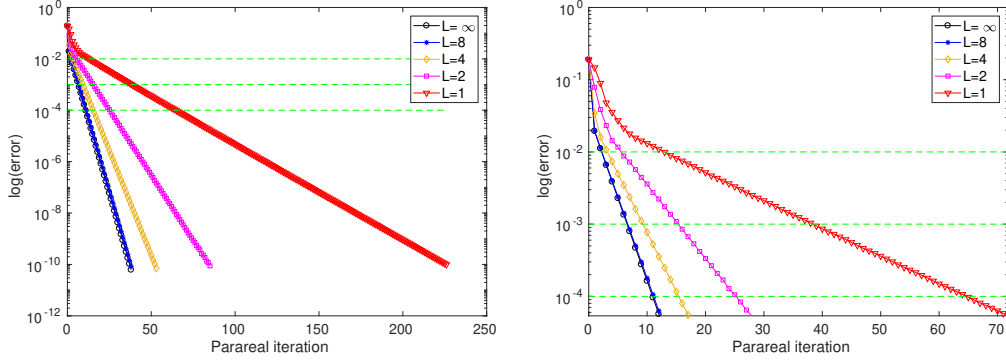


Figure 4: Example 2 (case 1): Relative error versus Parareal iterations, with a zoom on the first iterations on the right

L	1	2	4	8	10 ( $\approx \infty$ )	L	1	2	4	8	10 ( $\approx \infty$ )
$k$	13	6	4	3	3	$k$	39	16	10	7	7
$L * k$	13	12	16	24	30	$L * k$	39	32	40	56	70

L	1	2	4	8	10 ( $\approx \infty$ )
$k$	65	26	16	11	11
$L * k$	65	52	64	88	110

Table 5: Example 2 (case 1): Number of Parareal iterations  $k$  and total number of OSWRG iterations  $L * k$ , versus  $L$  to reach 3 different relative accuracy values:  $10^{-2}$  (top left),  $10^{-3}$  (top right) and  $10^{-4}$  (bottom)

Solver	OSWRG	Parareal	Parareal-OSWRG ( $L = 2$ )
iterations ( $n$ )	36	70	32
loss factor ( $\ell = n/36$ )	–	1.94	0.89
gain factor ( $N$ )	–	51	51
final gain factor ( $N/\ell$ )	–	26.29	57.38

Table 6: Example 2 (case 1): Gain factor for Parareal and coupled Parareal-OSWRG methods compared to the OSWRG solver for solving problem (43)

The time step of the fine solver is  $\Delta t_F = 0.01$ . In Figure 5, we plot the evolution of the relative error in  $L^\infty(0, T; H^1(\Omega))$ -norm, between the Parareal-OSWRG solution and the converged Parareal solution, as a function of the number of Parareal iterations, for different values of  $L$ . The Parareal algorithm ( $L = \infty$ ) corresponds to the case  $L = 12$ , and we will compare the other cases of  $L$  to this case. As for Case 1, the horizontal dashed green lines represent three possible values of 10% of the error, and Table 7 shows the number of Parareal and total OSWRG iterations to reach these different values, for  $L = 1, 2, 4, 8, 12$ . We observe that  $L = 2$  is the case with the fewest iterations.

The OSWRG algorithm used as a solver for problem (43) will need 24, 40 and 56 iterations to get a relative error in  $L^\infty(0, T; H^1(\Omega))$ -norm (between the iterative solution and the converged OSWRG solution) smaller than  $10^{-2}$ ,  $10^{-3}$  and  $10^{-4}$ , respectively. Table 8 shows the gain, when the fine solvers are performed in parallel, of the Parareal or the coupled Parareal-OSWRG methods compared to the OSWRG algorithm. In this table we take the values obtained for  $e = 10^{-2}$  (note that these values correspond also to those obtained with a mean value of the values for  $e = 10^{-1}$ ,  $e = 10^{-2}$  and  $e = 10^{-3}$ ). We observe that



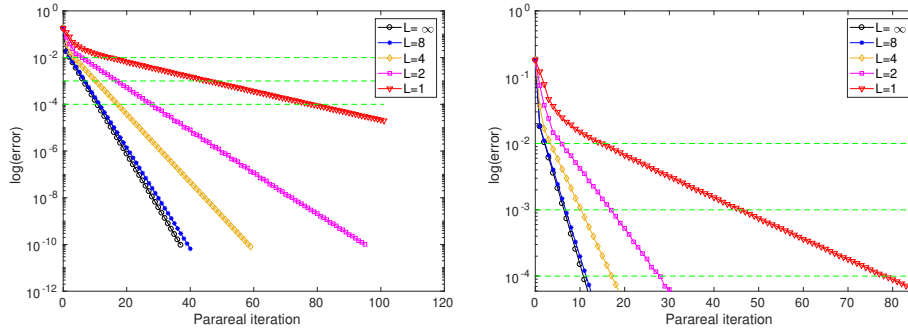


Figure 5: Example 2 (case 2): Relative error versus Parareal iterations, with a zoom on the first iterations on the right

for  $N = 51$  time windows, the results with a conforming mesh (Table 6) or with nonconforming meshes (Table 8) are very close and slightly better for nonconforming meshes adapted to the physics. Indeed, the gain is approximately a factor 24.28 for Parareal and a factor 60 for Parareal-OSWRG which is significantly better than the expected parallel efficiency.

L	1	2	4	8	12 ( $\approx \infty$ )	L	1	2	4	8	12 ( $\approx \infty$ )
$k$	16	6	4	3	3	$k$	47	17	11	7	
$L * k$	16	12	16	24	36	$L * k$	47	34	44	56	84

L	1	2	4	8	12 ( $\approx \infty$ )
$k$	79	28	17	11	
$L * k$	79	56	68	88	132

Table 7: Example 2 (case 2): Number of Parareal iterations  $k$  and total number of OSWRG iterations  $L * k$ , versus  $L$  to reach 3 different values:  $10^{-2}$  (top left),  $10^{-3}$  (top right) and  $10^{-4}$  (bottom)

Solver	OSWRG	Parareal	Parareal-OSWRG ( $L = 2$ )
iterations ( $n$ )	40	84	34
loss factor ( $\ell = n/40$ )	–	2.1	0.85
gain factor ( $N$ )	–	51	51
final gain factor ( $N/\ell$ )	–	24.28	60

Table 8: Example 2 (case 2): Gain factor for Parareal and coupled Parareal-OSWRG methods compared to the OSWRG solver for solving problem (43)

**Remark 7.1.** Note that in the results above, the number of OSWRG iterations to reach full convergence, inside each time window in the Parareal iterations, is smaller than the one needed to reach full convergence on  $(0, T)$  with OSWR method as a solver, probably because the time windows are smaller.

**Remark 7.2.** Another coupled algorithm is proposed in [5], involving incomplete (only one) iteration of OSWR both for the coarse and fine solvers (instead of incomplete iteration for the fine solver and a converged coarse solver as in our algorithm). The 2D simulations of this section don't show a comparison with that algorithm, however we have observed that the case  $L = 1$  is not the optimal choice of  $L$ , and that doing also incomplete iterations for the coarse solver will lead to similar or slower convergence than using a converged coarse solver.

### 7.3 Example in an industrial context

We consider a simplified model problem given by ANDRA, the French National Agency for Radioactive Waste Management (see also [18, 1, 2]), that simulates the transport of a contaminant in and around a nuclear waste repository site. The simulation domain is depicted in Figure 6 (left) (not at scale). The nuclear waste is stored in the repository (yellow), which is a 2950m by 10m rectangle located in the center of a clay domain of 3950m by 140m (light brown). In this example, we are concerned with the following time-dependent diffusion problem in mixed formulation:

$$\boldsymbol{\sigma} = -\mathbf{S}\nabla u \quad \text{in } \Omega \times (0, T), \quad (45a)$$

$$\phi \frac{\partial u}{\partial t} + \nabla \cdot \boldsymbol{\sigma} = f \quad \text{in } \Omega \times (0, T), \quad (45b)$$

where  $\Omega = [0, 3950] \times [0, 140]$ ,  $u$  represents the (dimensionless) concentration of the contaminant,  $\phi$  is the porosity, and  $\mathbf{S}$  is the time-independent diffusion tensor. We decompose  $\Omega$  into nine subdomains where  $\Omega_5$  is the nuclear waste repository domain, see Figure 6 (right). The porosity  $\phi$  and the tensor  $\mathbf{S}$  are defined as follows:

$$\phi = \begin{cases} 0.2 & \text{in } \Omega_5, \\ 0.05 & \text{in } \Omega_i, \quad i \neq 5, \end{cases} \quad \mathbf{S} = \begin{cases} 2 \times 10^{-9} \mathbb{I} \text{ m}^2/\text{s} & \text{in } \Omega_5, \\ 5 \times 10^{-12} \mathbb{I} \text{ m}^2/\text{s} & \text{in } \Omega_i, \quad i \neq 5, \end{cases}$$

where  $\mathbb{I}$  is the identity matrix. The source term  $f$  is zero in the clay layer and

$$f = \begin{cases} 10^{-5} \text{ years}^{-1} & \text{if } t \leq 10^5 \text{ years,} \\ 0 & \text{if } t > 10^5 \text{ years,} \end{cases} \quad \text{in the repository.}$$

The initial condition is  $u_0 = 0$  and we set homogeneous Dirichlet conditions on the top and the bottom of  $\Omega$ , and homogeneous Neumann conditions on the other sides of  $\partial\Omega$ . We are interested in the long-term behavior of the repository, over one million years, so that we set  $T = 10^6$  years. A dimensionless form of this problem is given in [2]. The number of triangles in the mesh of  $\Omega$  is 106638 (see Figure 7).

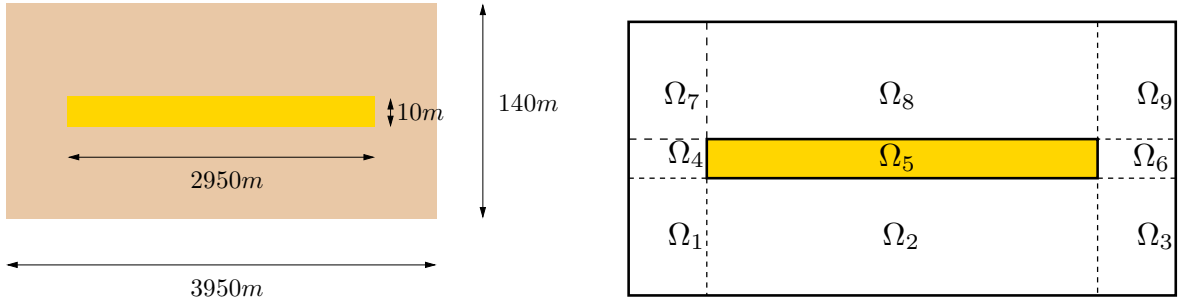


Figure 6: Geometry of the nuclear waste repository (yellow) and the clay layer around it (light brown) on the left and its decomposition into 9 subdomains on the right

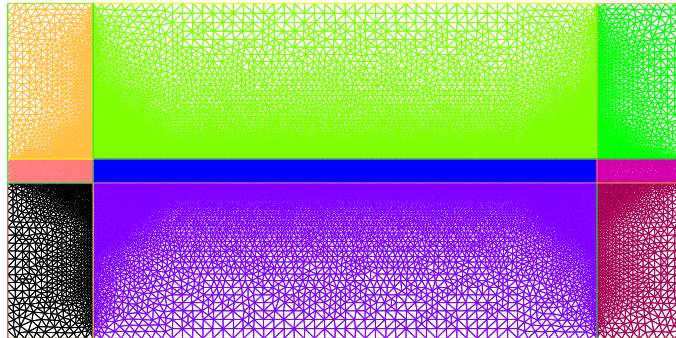


Figure 7: Example of a mesh used in and around a nuclear waste repository site

In Figure 8, we plot the evolution of the relative error in the  $L^2(\Omega \times (0, T))$ -norm, between the Parareal-OSWRG solution and the converged Parareal solution, as a function of the number of Parareal iterations, for different values of  $L$ . The Parareal algorithm ( $L = \infty$ ) corresponds to  $L = 10$  and we will compare the other cases of  $L$  to this case.

In [1], an a posteriori stopping criterion is given for this test case and the dashed green line shown in Figure 8 corresponds to this criterion. This line is reached after 3 iterations for Parareal and Parareal-OSWRG with  $L = 4$ , and after 12 iterations for Parareal-OSWRG with  $L = 2$ . The Parareal-OSWRG algorithm with  $L = 1$  seems to be not convergent (at least on the first thirty iterations). These data are reported in Table 9 and we observe that the fastest case is  $L = 4$  with a total of 12 OSWRG iterations.

Table 10 shows the gain, when the fine solvers are performed in parallel, of the Parareal or the coupled Parareal-OSWRG methods compared to the OSWRG algorithm. In the first column we give the number of OSWRG iterations given in [1] from the a posteriori stopping criterion, when the OSWRG method is used for solving problem (45). We observe that for  $N = 10$  time windows, we gain a factor 3.66 for Parareal and a factor 9.17 for Parareal-OSWRG which is almost the expected parallel efficiency.

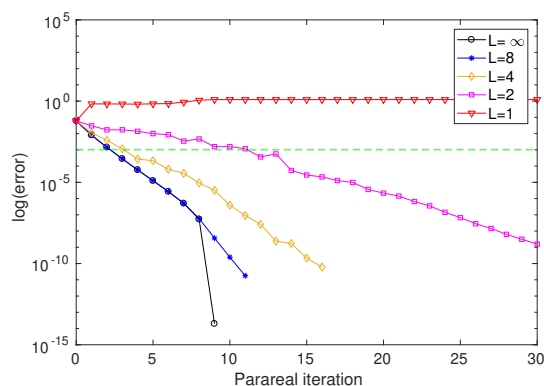


Figure 8: Example 3: Relative error versus Parareal iterations

L	1	2	4	8	10 ( $\approx \infty$ )
$k$	–	12	3	3	3
$L * k$	–	24	12	24	30

Table 9: Example 3: Number of Parareal iterations  $k$  and total number of OSWRG iterations  $L * k$ , versus  $L$

Solver	OSWRG	Parareal	Parareal-OSWRG ( $L = 4$ )
iterations ( $n$ )	11	30	12
loss factor ( $\ell = n/11$ )	–	2.73	1.09
gain factor ( $N$ )	–	10	10
final gain factor ( $N/\ell$ )	–	3.66	9.17

Table 10: Example 3: Gain factor for Parareal and coupled Parareal-OSWRG methods compared to the OSWRG solver for solving problem (45)

## A Proof of Theorem 4.2

*Proof.* The proof is done in [8] (Theorem 5.15) and we recall it here : for  $\ell = 1, 2, \dots, L$ , the error  $e_i^\ell := u_i^\ell - u|_{\Omega_i}$  satisfies

$$\mathcal{L}(e_i^\ell) = 0 \quad \text{in } \Omega_i \times \mathcal{I}, \quad (46a)$$

$$e_i^\ell(\cdot, 0) = 0 \quad \text{in } \Omega_i, \quad (46b)$$

$$\mathcal{B}_i e_i^\ell(0, \cdot) = \zeta_i^{\ell-1} \quad \text{in } \mathcal{I}, \quad (46c)$$

where

$$\zeta_i^\ell = \mathcal{B}_i e_i^\ell(0, \cdot), \quad \ell \geq 1. \quad (47)$$

Multiplying equation (46a) by  $e_i^\ell$  and integrating in space on  $\Omega_i$ , we obtain

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|e_1^\ell\|^2 + \nu \|\partial_x e_1^\ell\|^2 + b \|e_1^\ell\|^2 - (\nu \partial_x e_1^\ell - \frac{a}{2} e_1^\ell)(0, \cdot) e_1^\ell(0, \cdot) &= 0, \\ \frac{1}{2} \frac{d}{dt} \|e_2^\ell\|^2 + \nu \|\partial_x e_2^\ell\|^2 + b \|e_2^\ell\|^2 + (\nu \partial_x e_2^\ell - \frac{a}{2} e_2^\ell)(0, \cdot) e_2^\ell(0, \cdot) &= 0. \end{aligned}$$

Introducing the Robin interface operators  $\mathcal{B}_i$ ,  $i = 1, 2$ , defined in (2), and rewriting the terms on the interface in the form

$$(\nu \partial_x e_i^\ell - \frac{a}{2} e_i^\ell) e_i^\ell = \frac{1}{2p} ((\mathcal{B}_1 e_i^\ell)^2 - (\mathcal{B}_2 e_i^\ell)^2), \quad i = 1, 2,$$

we obtain the energy estimates, for  $i = 1, 2$

$$\frac{1}{2} \frac{d}{dt} \|e_i^\ell\|^2 + \nu \|\partial_x e_i^\ell\|^2 + b \|e_i^\ell\|^2 + \frac{1}{2p} ((\mathcal{B}_j e_i^\ell)(0, \cdot))^2 = \frac{1}{2p} ((\mathcal{B}_i e_i^\ell)(0, \cdot))^2, \quad j = 3 - i.$$

Replacing  $(\mathcal{B}_i e_i^\ell)(0, \cdot) = \zeta_i^{\ell-1}$ ,  $(\mathcal{B}_j e_i^\ell)(0, \cdot) = \zeta_j^\ell$  (from (46c) and (47)), and summing the above expression over  $i$ , we get, for all  $t \in \mathcal{I}$

$$\sum_i \left( \frac{1}{2} \frac{d}{dt} \|e_i^\ell\|^2 + \nu \|\partial_x e_i^\ell\|^2 + b \|e_i^\ell\|^2 \right) + \frac{1}{2p} \sum_i (\zeta_i^\ell)^2 = \frac{1}{2p} \sum_i (\zeta_i^{\ell-1})^2.$$

Then, integrating on  $\mathcal{I}$ , we obtain

$$\sum_i \left( \frac{1}{2} (\|e_i^\ell(\cdot, T)\|^2 - \|e_i^\ell(\cdot, 0)\|^2) + \int_{\mathcal{I}} (\nu \|\partial_x e_i^\ell\|^2 + b \|e_i^\ell\|^2) dt \right) + \frac{1}{2p} \|\zeta^\ell\|_{\mathcal{I}}^2 = \frac{1}{2p} \|\zeta^{\ell-1}\|_{\mathcal{I}}^2.$$

In  $L^2(0, T; H^1(\Omega_i))$ , we consider the norm

$$\|u\|_{L^2(0, T; H^1(\Omega_i))}^2 := \int_{\mathcal{I}} (\nu \|\partial_x u\|^2 + b \|u\|^2) dt.$$

Then we can rewrite the above equation as

$$\sum_i \left( \frac{1}{2} (\|e_i^\ell(\cdot, T)\|^2 - \|e_i^\ell(\cdot, 0)\|^2) + \|e_i^\ell\|_{L^2(0, T; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\zeta^\ell\|_{\mathcal{I}}^2 = \frac{1}{2p} \|\zeta^{\ell-1}\|_{\mathcal{I}}^2,$$

or equivalently

$$\frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0, T; H^1(\Omega_i))}^2 + \frac{1}{2p} \|\zeta^\ell\|_{\mathcal{I}}^2 = \frac{1}{2} \sum_i \|e_i^\ell(\cdot, 0)\|^2 + \frac{1}{2p} \|\zeta^{\ell-1}\|_{\mathcal{I}}^2.$$

Summing over  $\ell$  from 1 to  $L$ , we get a telescopic sum on the interface and thus

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0, T; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\zeta^L\|_{\mathcal{I}}^2 = \frac{1}{2} \sum_i \|e_i^\ell(\cdot, 0)\|^2 + \frac{1}{2p} \|\zeta^0\|_{\mathcal{I}}^2. \quad (48)$$

Using that  $e_i^\ell(\cdot, 0) = 0$ ,  $\forall \ell \geq 1$ ,  $i = 1, 2$ , we obtain

$$\sum_{\ell=1}^L \left( \frac{1}{2} \|e^\ell(\cdot, T)\|^2 + \sum_i \|e_i^\ell\|_{L^2(0, T; H^1(\Omega_i))}^2 \right) + \frac{1}{2p} \|\zeta^L\|_{\mathcal{I}}^2 = \frac{1}{2p} \|\zeta^0\|_{\mathcal{I}}^2.$$

Hence, the sum of the energies over all the iterates remains bounded, which implies that the energy in the iterates tends to zero when  $\ell \rightarrow \infty$ , which proves Theorem 4.2.  $\square$

## References

- [1] S. ALI HASSAN, *A posteriori error estimates and stopping criteria for solvers using domain decomposition method and with local time stepping*, PhD thesis, PhD thesis, Université Pierre et Marie Curie, Sorbonne Université, 2017.
- [2] S. ALI HASSAN, C. JAPHET, AND M. VOHRALÍK, *A posteriori stopping criteria for space-time domain decomposition for the heat equation in mixed formulations*, *Electron. Trans. Numer. Anal.*, 49 (2018), pp. 151–181.
- [3] G. BAL AND Y. MADAY, *A “parareal” time discretization for non-linear pde’s with application to the pricing of an american put*, in *Recent Developments in Domain Decomposition Methods*, L. F. Pavarino and A. Toselli, eds., Berlin, Heidelberg, 2002, Springer Berlin Heidelberg, pp. 189–202.
- [4] V. DOLEAN, P. JOLIVET, AND F. NATAF, *An introduction to domain decomposition methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015. Algorithms, theory, and parallel implementation.
- [5] M. GANDER, Y. JIANG, AND B. SONG, *A superlinear convergence estimate for the parareal Schwarz waveform relaxation algorithm*, *SIAM Journal on Scientific Computing*, 41 (2019), pp. A1148–A1169.
- [6] M. J. GANDER, *Optimized Schwarz Methods*, *SIAM J. Numerical Analysis*, 44 (2006), pp. 699–731.
- [7] M. J. GANDER AND E. HAIRER, *Nonlinear convergence analysis for the parareal algorithm*, in *Domain Decomposition Methods in Science and Engineering XVII*, U. Langer, M. Discacciati, D. E. Keyes, O. B. Widlund, and W. Zulehner, eds., Berlin, Heidelberg, 2008, Springer Berlin Heidelberg, pp. 45–56.
- [8] M. J. GANDER AND L. HALPERN, *Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems*, *SIAM J. Numer. Anal.*, 45 (2007), pp. 666–697.
- [9] M. J. GANDER, L. HALPERN, AND F. NATAF, *Optimal Convergence for Overlapping and Non-Overlapping Schwarz Waveform Relaxation*, in *Proceedings of the 11th International Conference on Domain Decomposition Methods*, C.-H. Lai, P. Bjørstad, M. Cross, and O. Widlund, eds., 1999, pp. 27–36.
- [10] ———, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, *SIAM J. Numer. Anal.*, 41 (2003), pp. 1643–1681.
- [11] ———, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, *SIAM J. Numer. Anal.*, 41 (2003), pp. 1643–1681.
- [12] M. J. GANDER, Y.-L. JIANG, AND R.-J. LI, *Parareal Schwarz waveform relaxation methods*, in *Domain Decomposition Methods in Science and Engineering XX*, R. Bank, M. Holst, O. Widlund, and J. Xu, eds., Berlin, Heidelberg, 2013, Springer Berlin Heidelberg, pp. 451–458.
- [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 S. VANDEWALLE, *Analysis of the parareal time-parallel time-integration method*, *SIAM J. Scientific Computing*, 29 (2007), pp. 556–578.
- [15] E. GILADI AND H. KELLER, *Space-time domain decomposition for parabolic problems*, *Numerische Mathematik*, 93 (2002), pp. 279–313.
- [16] R. GUETAT, *Méthode de parallélisation en temps : application aux méthodes de décomposition de domaine*, PhD thesis, Thèse de doctorat en Mathématiques appliquées, UPMC Université Paris 6 et Ecole polytechnique de Tunisie, 2011.
- [17] L. HALPERN, C. JAPHET, AND J. SZEFTTEL, *Optimized Schwarz waveform relaxation and discontinuous Galerkin time stepping for heterogeneous problems*, *SIAM J. Numer. Anal.*, 50 (2012), pp. 2588–2611.

- [18] T.-T.-P. HOANG, J. JAFFRÉ, C. JAPHET, M. KERN, AND J. E. ROBERTS, *Space-time domain decomposition methods for diffusion problems in mixed formulations*, SIAM J. Numer. Anal., 51 (2013), pp. 3532–3559.
- [19] T.-T.-P. HOANG, JAPHET, M. KERN, AND J. E. ROBERTS, *Space-Time Domain Decomposition for Reduced Fracture Models in Mixed Formulation*, SIAM J. Numer. Anal., 54 (2016), pp. 288–316.
- [20] C. JAPHET, Y. MADAY, AND F. NATAF, *A new interface cement equilibrated mortar method with Ventcel conditions*, in Domain decomposition methods in science and engineering XXI, J. Erhel, M. Gander, L. Halpern, G. Pichot, T. Sassi, and O. Widlund, eds., vol. 98 of Lect. Notes Comput. Sci. Eng., Springer, 2014, pp. 329–336.
- [21] C. JAPHET AND F. NATAF, *The best interface conditions for domain decomposition methods: absorbing boundary conditions*, in Absorbing Boundaries and Layers, Domain Decomposition Methods, Nova Sci. Publ., Huntington, NY, 2001, pp. 348–373.
- [22] C. JAPHET, F. NATAF, AND F. ROGIER, *The optimized order 2 method: Application to convection-diffusion problems*, Future Generation Computer Systems, 18 (2001), pp. 17–30.
- [23] J.-L. LIONS, Y. MADAY, AND G. TURINICI, *A "parareal" in time discretization of PDE's*, Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, 332 (2001), pp. 661–668.
- [24] J.-L. LIONS AND E. MAGENES, *Non-homogeneous boundary value problems and applications. Vol. I*, Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
- [25] ———, *Non-homogeneous boundary value problems and applications. Vol. II*, Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 182.
- [26] P.-L. LIONS, *On the Schwarz alternating method. III: a variant for nonoverlapping subdomains*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, held in Houston, Texas, March 20-22, 1989, R. G. J. P. T. F. Chan and O. Widlund, eds., Philadelphia, PA, SIAM, 1990, pp. 202–223.
- [27] Y. MADAY AND O. MULA, *An adaptive parareal algorithm*, Journal of Computational and Applied Mathematics, 377 (2020), p. 112915.
- [28] Y. MADAY AND G. TURINICI, *The parareal in time iterative solver: a further direction to parallel implementation*, in Domain Decomposition Methods in Science and Engineering, T. J. Barth, M. Griebel, D. E. Keyes, R. M. Nieminen, D. Roose, T. Schlick, R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. Widlund, and J. Xu, eds., Berlin, Heidelberg, 2005, Springer Berlin Heidelberg, pp. 441–448.
- [29] V. MARTIN, *An optimized Schwarz waveform relaxation method for the unsteady convection diffusion equation in two dimensions*, Appl. Numer. Math., 52 (2005), pp. 401–428.
- [30] O. MULA HERNANDEZ, *Quelques contributions vers la simulation parallèle de la cinétique neutronique et la prise en compte de données observées en temps réel*, PhD thesis, Thèse de doctorat dirigée par Maday, Yvon Mathématiques appliquées Paris 6 2014, 2014.
- [31] P. SMOLARKIEWICZ, *The multi-dimensional crowley advection scheme*, Monthly Weather Review, 110 (1982), pp. 1968–1983.