Journal of Computational Physics ••• (••••) ••••••

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Convergence analysis and optimization of a Robin Schwarz waveform relaxation method for time-periodic parabolic optimal control problems

Gabriele Ciaramella [a,*], Laurence Halpern [b], Luca Mechelli [c]

[a] MOX, Dipartimento di Matematica, Politecnico di Milano, Milan, Italy
[b] Université Sorbonne Paris Nord, Paris, France
[c] Universität Konstanz, Constance, Germany

A R T I C L E   I N F O

Keywords:
Domain decomposition methods
Schwarz methods
Optimal control problems
Periodic parabolic equations
Discrete Fourier analysis

A B S T R A C T

This paper is concerned with a novel convergence analysis of the optimized Schwarz waveform relaxation method (OSWRM) for the solution of optimal control problems governed by periodic parabolic partial differential equations (PDEs). The new analysis is based on a Fourier-type technique applied to a semidiscrete-in-time form of the optimality condition. This leads to a precise characterization of the convergence factor of the method at the semidiscrete level. Using this characterization, the optimal transmission condition parameter is obtained at the semidiscrete level and its asymptotic behavior as the time discretization converges to zero is analyzed in detail.

1. Introduction

The control of time-periodic PDEs plays an important role in several applications, like the control of eddy current electromagnetic problems [1–4] and Stokes problems [5], energy-producing kites [6], cyclically steered bio-reactors [7], design of reverse flow reactors [8], control of magnetohydrodynamic phenomena [9,10] and related multiharmonic models [11]. In this scenario, time-periodic parabolic problems are considered in [5,6,8]. For this important class of problems different solvers and preconditioners, like finite-element solvers, multigrid methods, and algebraic preconditioners have been developed and analyzed; see, e.g., [12–15] and references therein. Also domain decomposition methods have been used for PDE-constrained optimization problems [16,17]. For elliptic optimal control problems classical Schwarz methods were considered in [18] as preconditioners (see also [19]), while in [19–22], optimized Schwarz methods have been introduced and analyzed. Neumann-Neumann methods are studied in [23,24]. Robin Schwarz waveform relaxation methods were introduced in [25]. OSWRMs are Schwarz domain decomposition methods characterized by Robin transmission conditions, where the choice of the Robin-type parameter affects tremendously the convergence of the method; see, e.g., [26–29]. In the context of parabolic control problems, the only convergence analysis proposed in the literature (and that can be adapted to time-periodic problems) is the one presented in [25] and based on energy estimates. However, this analysis does not lead to a concrete estimate of the convergence factor and does not provide insights that can be used to choose the Robin parameter.

The goal of this paper is to present a novel Fourier-type convergence analysis of an OSWRM for the solution of optimal control problems governed by time-periodic parabolic equations. In particular, we perform a semidiscrete-in-time analysis that allows us to obtain precise estimates of the convergence factor, which can be used to optimize the Robin parameter characterizing the transmission

* Corresponding author.
E-mail addresses: gabriele.ciaramella@polimi.it (G. Ciaramella), halpern@math.univ-paris13.fr (L. Halpern), luca.mechelli@uni-konstanz.de (L. Mechelli).

https://doi.org/10.1016/j.jcp.2023.112572
Received 6 September 2022; Received in revised form 12 June 2023; Accepted 13 October 2023

conditions. Although we could perform a continuous Fourier analysis, we carried out a semidiscrete one in time, since it gives a better characterization of the numerical behavior of the OSWRM (see [30]). Moreover, our analysis permits us to obtain a convergence result not only in the nonoverlapping case (for which convergence can be proved by energy estimates [20]), but also in the overlapping case. In particular, in the semidiscrete case convergence of nonoverlapping methods is guaranteed by the compactness of the set of possible Fourier frequencies. Thus, one can prove that the contraction factor is smaller than a constant lower than one. This is not possible in the continuous setting, for which the set of Fourier frequencies is unbounded. In this case, Parseval's identity together with the dominated convergence (Lebesgue) theorem need to be used. However, our optimization study concerns both cases.

The optimal Robin parameter for the semidiscrete case is obtained by solving an inf-sup problem. Similar problems have been treated in the literature, see, e.g., [19,27,28,31]. Although we will use few of the results contained in these works, there are three main differences in our contribution: the inf-sup problem is defined on the Cassini ovals (cf. Remark 2.3), the problem is not convex, and furthermore the Robin parameter is constrained to be real. We point out that, even though the proposed analysis is carried out for a one-dimensional space domain, its development is already very involved. The two-dimensional case could be explored by using a similar technique, but it will require further attention to details complicating the presentation of the results.

The paper is organized as follows. The optimal control problem and the OSWRM are introduced in Section 2. In Section 3, convergence of the OSWRM is proved and its convergence factor is characterized in terms of the parameters of the problem. The optimal parameter p is computed in Section 4 in the case of non-overlapping and overlapping subdomains. In particular, while in the nonoverlapping case we are able to obtain a precise formula for the optimal parameter, this is not possible in the overlapping case, where asymptotic expressions are instead derived. We will distinguish two cases depending on the relation between the overlap L and the size of the time grid Δt. First, the overlap L is chosen proportionally to Δt. Second, L is chosen proportionally to $\sqrt{\Delta t}$. In Section 5, we demonstrate the validity of our theoretical findings by direct numerical experiments. Finally, we outline our conclusion in Section 6.

2. The optimal control problem and the OSWRM

Let $\Omega = \mathbb{R}$ denote the spatial domain and $[0, T]$ the time domain, with the final time $T > 0$. Consider the quadratic cost functional

$$J(y,u) := \frac{1}{2}\|y - y_Q\|^2_{L^2((0,T)\times\Omega)} + \frac{\sigma}{2}\|u\|^2_{L^2((0,T)\times\Omega)}, \tag{1}$$

where σ is a positive parameter which penalizes the size of the control u, and y_Q is a target state. The state y is subject to the linear parabolic constraint

$$\begin{aligned} \partial_t y - \lambda\partial_{xx} y + d\,y = u, &\qquad \text{in } (0,T)\times\Omega, \\ y(0) = y(T), &\qquad \text{in } \Omega, \end{aligned} \tag{2}$$

with $\lambda, d > 0$. The target state y_Q is in $L^2((0,T)\times\Omega)$. For any $u \in L^2((0,T)\times\Omega)$, Problem (2) has a unique solution $y \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$, the regularity theorems imply that $y \in C([0,T]\times\Omega)$, which justifies the boundary condition in time, see [32]. Furthermore, y is a linear function of u. The σ-convexity of the quadratic map $u \mapsto J(y(u),u)$ implies existence and uniqueness of the minimizer $(y,u) \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega)) \times L^2((0,T)\times\Omega)$, see [33]. Moreover, the unique minimizer (y,u) is characterized by the first-order optimality system consisting of (2), completed by the adjoint equation [34,35]

$$\begin{aligned} -\partial_t q - \lambda\partial_{xx} q + d\,q = y_Q - y &\qquad \text{in } (0,T)\times\Omega, \\ q(T) = q(0), &\qquad \text{in } \Omega, \end{aligned} \tag{3}$$

and the condition

$$\sigma u - q = 0 \quad \text{in } (0,T)\times\Omega. \tag{4}$$

For a given y, the backward parabolic problem with final time condition (3) has similarly a unique solution $q \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$.

Now, let us introduce the OSWRM [27,36] for the solution of (2), (3), (4), written in a substructured form on the interface. We consider the decomposition $\Omega = \Omega_1 \cup \Omega_2$ with $\Omega_1 = (-\infty, x_1)$ and $\Omega_2 = (x_2, +\infty)$, with $x_1 - x_2 = L \geq 0$. For positive p, the iteration of the Schwarz waveform relaxation algorithm is defined by the $\mathcal{T}$ operator:

$$\mathcal{T}(\underline{g}_1, \underline{g}_2) = (\underline{g}_1', \underline{g}_2') :$$

$$\text{For } j = 1,2, \tag{5a}$$

$$\text{Given } \underline{g}_j = (g_j, \tilde{g}_j), \text{solve the forward-backward problem:}$$

$$q_j = \sigma u_j, \tag{5b}$$

$$\left| \begin{aligned} &\partial_t y_j - \lambda\partial_{xx} y_j + d\,y_j = u_j \text{ in } (0,T)\times\Omega_j, \\ &\partial_{\boldsymbol{n}_j} y_j(\cdot, x_j) + p y_j(\cdot, x_j) = g_j \text{ in } (0,T), \\ &y_j(0) = y_j(T) \text{ in } \Omega_j, \end{aligned} \right. \tag{5c}$$

$$\begin{cases} -\partial_t q_j - \lambda\partial_{xx} q_j + d q_j = y_Q - y_j \text{ in } (0,T)\times\Omega_j, \\ \partial_{\boldsymbol{n}_j} q_j(\cdot,x_j) + p q_j(\cdot,x_j) = \tilde{g}_j \text{ in } (0,T), \\ q_j(T) = q_j(0), \text{ in } \Omega_j. \end{cases} \tag{5d}$$

Compute for $i \neq j$

$$\begin{cases} g_i' = \partial_{\boldsymbol{n}_i} y_j(\cdot,x_i) + p y_j(\cdot,x_i) \text{ in } (0,T), \\ \tilde{g}_i' = \partial_{\boldsymbol{n}_i} q_j(\cdot,x_i) + p q_j(\cdot,x_i) \text{ in } (0,T), \\ \underline{g}_i = (g_i', \tilde{g}_i'). \end{cases} \tag{5e}$$

Here, $\partial_{\boldsymbol{n}_j}$ is the outward normal derivative at point x_j for $j=1,2$. The parameter $p>0$ is used to define Robin transmission conditions and its choice strongly influences the convergence of the method [27,36,37].

For a proper definition of the overlapping algorithm with the heat equation, we need more regularity, and use the anisotropic Sobolev spaces [38]

$$H^{2r,r}(O\times(0,T)) := L^2(0,T;H^{2r}(O)) \cap H^r(0,T;L^2(O)),$$

where O is a generic domain. We will mainly use $H^{3,\frac{3}{2}}$. Any u in this space has traces at $t=0$ and $t=T$, which belong to $H^2(O)$, and traces on the boundary of O, $\gamma_0 u \in H^{\frac{5}{4}}(0,T), \gamma_1 u = \frac{\partial u}{\partial n} \in H^{\frac{3}{4}}(0,T)$. For $r>\frac{1}{2}$, define the periodic space $H^r_{\#}(0,T)$ to be the space of functions in $H^r(0,T)$ (therefore continuous) which coincide at 0 and T. The application

$$u \mapsto (\gamma_0 u, \gamma_1 u): \quad H^{3,\frac{3}{2}}_{\#}((0,T)\times O) \to H^{\frac{5}{4}}_{\#}(0,T)\times H^{\frac{3}{4}}_{\#}(0,T)$$

is linear continuous and surjective. This is an extension of results in [32, Theorem 2.3, p. 21] and replaces the usual compatibility conditions between the trace and the initial condition. Let $\mathsf{S}_j(\cdot;g): L^2((0,T)\times\Omega_j) \to H^{2,1}_{1,\#}((0,T)\times\Omega_j)$ be the solution operator associated to the state equation (5c), that is $\mathsf{S}(u_j;\underline{g}) = y_j$. Here, $\underline{g}$ represents a boundary data or a source term. We can now prove the following result. In what follows, we denote by $(\cdot,\cdot)$ the usual inner product for $L^2((0,T)\times\Omega)$ and by $(\cdot,\cdot)_{L^2(0,T)}$ the inner product for $L^2(0,T)$.

Lemma 1. *In each subdomain, for a given* $\underline{g}_j = (g_j, \tilde{g}_j) \in (L^2(0,T))^2$, *define the cost function*

$$J_j(y_j,u_j) = \frac{1}{2}\|y_j - y_Q\|^2_{L^2((0,T)\times\Omega_j)} + \frac{\sigma}{2}\|u_j\|^2_{L^2((0,T)\times\Omega_j)} - \lambda(\tilde{g}_j, y_j(\cdot,x_j))_{L^2(0,T)}, \tag{6}$$

where for u_j *in* $L^2((0,T)\times\Omega_j)$, y_j *is the solution to (5c). Moreover, define the reduced cost functional* $\widehat{J}_j(u_j) := J_j(\mathsf{S}_j(u_j;\underline{g}_j),u_j)$. $\widehat{J}_j(u_j)$ *is* σ*-convex on* $L^2((0,T)\times\Omega_j)$, *and the equations (5b), (5c), (5d) form the optimality system for the minimization of* $\widehat{J}_j$.

Proof. The proof is similar to those given in [21] and [39]. The cost functional $\widehat{J}_j$ is differentiable and σ-convex, it has one and only one minimum $\bar{u}_j$, characterized by $\widehat{J}_j'(\bar{u}_j) = 0$ [35]. Compute now the derivative of $\widehat{J}_j$:

$$\begin{aligned} \widehat{J}_j'(u_j)\cdot h =& (\mathcal{S}(u_j;\underline{g}_j) - y_Q, \mathcal{S}'(u_j;0)\cdot h) + \sigma(u_j,h) - \lambda(\tilde{g}_j, z(\cdot,x_j))_{L^2(0,T)} \\ =& (y_j - y_Q, z) + \sigma(u_j,h) - \lambda(\tilde{g}_j, z(\cdot,x_j))_{L^2(0,T)}, \end{aligned}$$

where $z = \mathcal{S}(h;0)$. In order to identify the quantity above as a scalar product, introduce q_j solution of (5d). We have then, by integration by parts,

$$\begin{aligned} (y_j - y_Q, z) &= (-\partial_t q_j - \lambda\partial_{xx} q_j + d q_j, z) \\ &= (\partial_t z - \lambda\partial_{xx} z + d z, q_j) - [(q_j(t,\cdot), z(t,\cdot))_{L^2(\Omega_j)}]_0^T \\ &\quad + \lambda\big(-(\partial_x q_j(x_j), z(x_j))_{L^2(0,T)} + (q_j(x_j), \partial_x z(x_j))_{L^2(0,T)}\big). \end{aligned}$$

Thanks to the periodicity conditions, the second term vanishes. Using the heat equation on z, the first one is equal to (q_j,h). As for the boundary terms, use the boundary conditions in the equations to get

$$-(\partial_x q_j(x_j), z(x_j))_{L^2(0,T)} + (q_j(x_j), \partial_x z(x_j))_{L^2(0,T)} = (\tilde{g}_j, z(\cdot,x_j))_{L^2(0,T)},$$

which cancels out with the boundary term in $\widehat{J}_j'(u_j)\cdot h$. There remains only

$$\widehat{J}_j'(u_j)\cdot h = (\sigma u_j + q_j, h).$$

Hence, $\widehat{J}_j'(u_j)$ can be identified with $\sigma u_j + q_j$. Thus, the last equation to identify the optimality system is $\sigma u_j + q_j = 0$. $\square$

Theorem 2 *(Well-posedness of the OSWRM). For any target state y_Q in $H^{1,\frac{1}{2}}((0,T)\times\Omega)$, the iteration map $\mathcal{T}$ defined by (5) maps $(H_\#^{3/4}(0,T))^4$ into $(H_\#^{3/4}(0,T))^4$. For any initialization $\underline{g}^0=(\underline{g}_1^0,\underline{g}_2^0)\in(H_\#^{\frac{3}{4}}((0,T))^4$, it defines a sequence $\underline{g}^n=(\underline{g}_1^n,\underline{g}_2^n)\in(H_\#^{\frac{3}{4}}((0,T))^4$ by the linear recursion $\underline{g}^n=\mathcal{T}\,\underline{g}^{n-1}$. Associated to $\underline{g}^n$ are $y^{j,n}$ and $q^{j,n}\in H_\#^{3,\frac{3}{2}}((0,T)\times\Omega_j)$ for $j=1,2$, defined by (5b), (5c), (5d).*

Proof. By Lemma 1, $\hat{J}_j(u_j)$ is a quadratic σ-convex function, and therefore has a unique minimum point for $\tilde{g}_j\in L^2(0,T)$, characterized by the optimality system, which is precisely (5). Now, if $\tilde{g}_j\in H^{\frac{3}{4}}(0,T)$, by the regularity results in [32, Theorem 2.1] cited above, y_j and q_j are in $H_\#^{3,\frac{3}{2}}((0,T)\times\Omega_j)$, and by the trace theorems at point x_i, g_i' and $\tilde{g}_i'$ are in $H^{\frac{3}{4}}(0,T)$. □

The convergence of the algorithm can be obtained through *a priori* estimates in the case of nonoverlapping subdomains [20,40]. In the overlapping case, the appropriate tool is Fourier series, using the periodicity of the problem. We do not carry out the computation, since it is very similar to the one we perform in the next section on the semi-discrete case. Similarly, we do not carry out the optimization of the Robin parameter, since it is reasonable to expect, and it was proven in the elliptic case, that a semi-discrete optimization is more relevant to the actual computations.

3. The semidiscrete algorithm

In this section, we carry out a convergence analysis for the semidiscrete-in-time domain decomposition algorithm. The semidiscrete systems are obtained using the implicit Euler scheme, as it is usual for parabolic equations. As in the continuous case [27,36,37,41], we identify the subproblems as control problems for a modified cost function, which permits to prove the well-posedness of the algorithm. The convergence is obtained through a discrete Fourier transform. A discrete Fourier analysis for stationary problems can be found in, e.g., [30,42].

3.1. Definition and well-posedness

Introduce a uniform grid of size $\Delta t=T/S$, that discretizes the interval $[0,T]$ with gridpoints $t_s=s\Delta t$ for $s=0,\dots,S$. The functions y and q of t and x are approximated by vectors Y and Q in $\mathbb{R}^{S+1}$, functions of x, with components indexed by s. Y_Q is the vector defined by $(Y_Q)_s=y_Q(t_s)$. We discretize the state equation (2) and the adjoint equation (3) in time by an implicit Euler scheme and obtain

$$\left|\begin{array}{l}\frac{1}{\Delta t}\left(Y_s-Y_{s-1}\right)-\lambda\partial_{xx}Y_s+dY_s=U_s \text{ in } [\![1,S]\!]\times\Omega,\\ Y_0=Y_S \text{ in } \Omega,\end{array}\right. \tag{7a}$$

$$\sigma U=Q \text{ in } \text{ in } [\![1,S]\!]\times\Omega, \tag{7b}$$

$$\left|\begin{array}{l}\frac{1}{\Delta t}\left(Q_s-Q_{s+1}\right)-\lambda\partial_{xx}Q_s+dQ_s=(Y_Q)_s-Y_s \text{ in } [\![0,S-1]\!]\times\Omega,\\ Q_S=Q_0 \text{ in } \Omega,\end{array}\right. \tag{7c}$$

where for M and N two integers, $[\![M,N]\!]$ denotes the set of integers between M and N (including M and N). We define $R_\#:=\{X\in(L^2(\Omega))^{S+1},\ X_0=X_S\}$.

Theorem 3. *For any $r\geq 0$ and $U\in(H^r(\Omega))^{S+1}\cap R_\#$, Problem (7a) has a unique solution $Y\in(H^{r+2}(\Omega))^{S+1}\cap R_\#$. Similarly, for any $Y,Y_Q\in(H^r(\Omega))^{S+1}\cap R_\#$, Problem (7c) has a unique solution $Q\in(H^{r+2}(\Omega))^{S+1}\cap R_\#$. Moreover, the system (7a)-(7c) has a unique solution $(Y,U,Q)\in(H^r(\Omega))^{S+1}\cap R_\#\times(H^r(\Omega))^{S+1}\times(H^{r+2}(\Omega))^{S+1}\cap R_\#$.*

Proof. We apply the discrete Fourier Transform (DFT) to $(Y_0,\dots,Y_{S-1})$. Given a vector $X=(X_0,\dots,X_{S-1})\in\mathbb{R}^S$, the DFT is given by $\widehat{X}=(\widehat{X}_0,\dots,\widehat{X}_{S-1})\in\mathbb{R}^S$, where $\widehat{X}_\kappa=\sum_{s=0}^{S-1}X_se^{-i2\pi\kappa s/S}$. The inverse DFT is then $X_s=\frac{1}{S}\sum_{\kappa=0}^{S-1}\widehat{X}_\kappa e^{i2\pi\kappa s/S}$ and the Parseval equality holds: $\sum_{s=0}^{S-1}|X_s|^2=\sum_{s=0}^{S-1}|\widehat{X}_s|^2$. Thus, (7a) becomes

$$d_S(\kappa)\widehat{Y}_\kappa-\lambda\partial_{xx}\widehat{Y}_\kappa=\widehat{U}_\kappa,\ \kappa\in[\![0,S-1]\!], \tag{8}$$

where $d_S(\kappa):=\left(\frac{S}{T}(1-e^{-\frac{2\pi i}{S}\kappa})+d\right)\in\mathbb{C}$. Since the real part of d_S is bounded from below by d, the problem above is strongly elliptic and has a unique solution with $\widehat{Y}_\kappa\in H^{r+2}(\Omega)$. The inverse DFT gives the existence and uniqueness of a solution to Problem (7a).

The proof applies to (7c) as well. Finally, notice that (7a)-(7c) is the first-order optimality system of a linear-quadratic and strictly convex optimal control problem (similar to Theorem 5) and hence uniquely solvable. □

We also discretize in time the iteration of the Schwarz algorithm (5):

$$\mathcal{T}_{\Delta t}(\underline{G}_1, \underline{G}_2) = (\underline{G}'_1, \underline{G}'_2):$$

For $j = 1, 2$ (9a)

Given $\underline{G}_j = (G_j, \tilde{G}_j) \in R^2_\#$, solve

$$Q_j = \sigma U_j \tag{9b}$$

$$\left|\begin{array}{l} \dfrac{Y_j(s) - Y_j(s-1)}{\Delta t} - \lambda \partial_{xx} Y_j(s) + d Y_j(s) = U_j(s) \text{ in } [\![1, S]\!] \times \Omega_j, \\ \partial_{\boldsymbol{n}_j} Y_j(\cdot, x_j) + p Y_j(\cdot, x_j) = G_j \text{ in } [\![0, S]\!], \\ Y_j(0, \cdot) = Y_j(S, \cdot) \text{ in } \Omega_j, \end{array}\right. \tag{9c}$$

$$\left|\begin{array}{l} \dfrac{Q_j(s) - Q_j(s+1)}{\Delta t} - \lambda \partial_{xx} Q_j(s) + d Q_j(s) = Y_Q(s) - Y_j(s) \text{ in } [\![0, S-1]\!] \times \Omega_j, \\ \partial_{\boldsymbol{n}_j} Q_j(\cdot, x_j) + p Q_j(\cdot, x_j) = \tilde{G}_j \text{ in } [\![0, S]\!], \\ Q_j(0, \cdot) = Q_j(S, \cdot) \text{ in } \Omega_j. \end{array}\right. \tag{9d}$$

Compute for $i \neq j$, in $[\![0, S]\!]$,

$$\left|\begin{array}{l} G'_i = \partial_{\boldsymbol{n}_i} Y_j(\cdot, x_i) + p Y_j(\cdot, x_i) \text{ in } [\![0, S]\!], \\ \tilde{G}'_i = \partial_{\boldsymbol{n}_i} Q_j(\cdot, x_i) + p Q_j(\cdot, x_i) \text{ in } [\![0, S]\!], \\ \underline{G}'_i = (G'_i, \tilde{G}'_i) \in R^2_\#. \end{array}\right. \tag{9e}$$

Lemma 4. *For any $r \geq 0$ and $U_j \in (H^r(\Omega_j))^{S+1} \cap R_\#$, for any $G_j \in R_\#$, Problem (9c) has a unique solution $Y_j \in (H^{r+2}(\Omega))^{S+1} \cap R_\#$. Similarly, for any $Y_j, Y_Q|_{\Omega_j} \in (H^r(\Omega_j))^{S+1} \cap R_\#$, $\tilde{G}_j \in R_\#$, Problem (9d) has a unique solution $Q_j \in (H^{r+2}(\Omega_j))^{S+1} \cap R_\#$.*

Proof. The proof goes by DFT, similar to that of the previous lemma. □

The spaces $\mathbb{R}^{S+1}$ and $\mathbb{L}^2(\Omega) = L^2(\Omega)^{S+1}$ are equipped with the norms

$$\|Y\|_S^2 = \Delta t \sum_{s=1}^{S} |Y_s|^2, \quad \|Y\|^2_{\mathbb{L}^2(\Omega)} = \Delta t \sum_{s=1}^{S} \|Y_s\|^2_{L^2(\Omega)}.$$

Theorem 5. *The system (9b)-(9d) is the optimality system for the minimization of*

$$J_j(U, Y) = \frac{1}{2}\|Y - Y_Q\|^2_{\mathbb{L}^2(\Omega_j)} + \frac{\sigma}{2}\|U\|^2_{\mathbb{L}^2(\Omega_j)} - \lambda(\tilde{G}_j, Y_j(x_j))_S, \tag{10}$$

subject to (9c). Therefore (9) defines a continuous linear operator $\mathcal{T}_{\Delta t}$, $R^4_\# \to R^4_\#$.

Proof. Thanks to Lemma 4, the minimization problem is well-defined. It is a quadratic σ-convex problem, thus has a single solution, characterized by the optimality system. The proof that the optimality system is (9b)-(9d) is analog to the proof in the continuous case in Lemma 1, replacing, for the time integration by parts, continuous by discrete. The detailed calculations are reported in the Appendix. □

The semidiscrete algorithm is now defined by

$$\underline{G}^0 \in R^4_\#, \quad \underline{G}^n = \mathcal{T}_{\Delta t} \underline{G}^{n-1} \in R^4_\#. \tag{11}$$

3.2. Semidiscrete convergence analysis

To study convergence of the semidiscrete OSWRM, we apply the iteration to the error $\mathcal{Y}_j = Y - Y_j$, $\mathcal{U}_j = U - U_j$ and $\mathcal{Q}_j = Q - Q_j$. Thus, denoting by $\underline{\mathcal{G}}_j$ the Robin traces in error form, and by $\underline{\widehat{\mathcal{G}}}_j$ the corresponding (discrete) Fourier transformed elements, we can introduce the discrete Fourier transformed system as in (8), that is

$$\widehat{\mathcal{T}}_{\Delta t}(\underline{\widehat{\mathcal{G}}}_1, \underline{\widehat{\mathcal{G}}}_2) = (\underline{\widehat{\mathcal{G}}}'_1, \underline{\widehat{\mathcal{G}}}'_2):$$

For $j = 1, 2$ (12a)

Given $\underline{\widehat{\mathcal{G}}}_j = (\widehat{\mathcal{G}}_j, \widehat{\tilde{\mathcal{G}}}_j)$, solve

$$\widehat{Q}_j = \sigma \widehat{U}_j, \tag{12b}$$

$$\left|\begin{array}{l} d_S\widehat{\mathcal{Y}}_j - \lambda\partial_{xx}\widehat{\mathcal{Y}}_j = \frac{1}{\sigma}\widehat{\mathcal{U}}_j \text{ in } [\![0,S]\!]\times\Omega_j, \\ \partial_{n_j}\widehat{\mathcal{Y}}_j(x_j) + p\widehat{\mathcal{Y}}_j(x_j) = \widehat{\mathcal{G}}_j \text{ in } [\![0,S]\!], \end{array}\right. \tag{12c}$$

$$\left|\begin{array}{l} \overline{d_S(\kappa)}\widehat{\mathcal{Q}}_j - \lambda\partial_{xx}\widehat{\mathcal{Q}}_j = -\widehat{\mathcal{Y}}_j \text{ in } [\![0,S]\!]\times\Omega_j, \\ \partial_{n_j}\widehat{\mathcal{Q}}_j(x_j) + p\widehat{\mathcal{Q}}_j(x_j) = \widehat{\widetilde{\mathcal{G}}}_j \text{ in } [\![0,S]\!]. \end{array}\right. \tag{12d}$$

Compute for $i \neq j$, in $[\![0,S]\!]$,

$$\left|\begin{array}{l} \widehat{\mathcal{G}}_i' = \partial_{n_i}\widehat{\mathcal{Y}}_j(x_i) + p\widehat{\mathcal{Y}}_j(x_i), \\ \widehat{\widetilde{\mathcal{G}}}_i' = \partial_{n_i}\widehat{\mathcal{Q}}_j(x_i) + p\widehat{\mathcal{Q}}_j(x_i), \\ \underline{\widehat{\mathcal{G}}}_i' = (\widehat{\mathcal{G}}_i', \widehat{\widetilde{\mathcal{G}}}_i'). \end{array}\right. \tag{12e}$$

The recursion gives the sequence of vectors on the interfaces

$$\underline{\widehat{\mathcal{G}}}^n = \widehat{\mathcal{T}_{\Delta t}}\underline{\widehat{\mathcal{G}}}^{n-1} = \widehat{\mathcal{T}_{\Delta t}}^{\,n}\underline{\widehat{\mathcal{G}}}^0. \tag{13}$$

The following lemma summarizes the notations we use in what follows and computes the iteration matrix $\widehat{\mathcal{T}_{\Delta t}}$.

Lemma 6. *Define*

$$\begin{aligned} &d_S(\kappa) := \frac{S}{T}(1 - e^{-\frac{2\pi i}{S}\kappa}) + d, \;\; \mu_S(\kappa) := \frac{1}{\lambda}\left(\mathrm{Re}\left(d_S\right) + i\sqrt{\frac{1}{\sigma} + \left(\mathrm{Im}(d_S)\right)^2}\right), \\ &P := \begin{bmatrix} i\left(\mathrm{Im}(d_S) + \sqrt{\left(\mathrm{Im}(d_S)\right)^2 + \frac{1}{\sigma}}\right) & i\left(\mathrm{Im}(d_S) - \sqrt{\left(\mathrm{Im}(d_S)\right)^2 + \frac{1}{\sigma}}\right) \\ 1 & 1 \end{bmatrix}, \\ &z_S := \sqrt{\mu_S}, \quad \rho_S(\kappa, p, L) = \frac{z_S(\kappa) - p}{z_S(\kappa) + p}e^{-z_S(\kappa)L}, \quad G_S := \begin{pmatrix} \rho_S & 0 \\ 0 & \overline{\rho_S} \end{pmatrix}. \end{aligned} \tag{14}$$

Then the iteration is explicitly given by

$$\begin{pmatrix} \underline{\widehat{\mathcal{G}}}_1' \\ \underline{\widehat{\mathcal{G}}}_2' \end{pmatrix} = \begin{pmatrix} 0 & P^{-1}G_S P \\ P^{-1}G_S P & 0 \end{pmatrix}\begin{pmatrix} \underline{\widehat{\mathcal{G}}}_1 \\ \underline{\widehat{\mathcal{G}}}_2 \end{pmatrix} =: \widehat{\mathcal{T}_{\Delta t}}\begin{pmatrix} \underline{\widehat{\mathcal{G}}}_1 \\ \underline{\widehat{\mathcal{G}}}_2 \end{pmatrix}. \tag{15}$$

Proof. Defining the vectors $X_j = (\widehat{Y}_j, \widehat{Q}_j)$ for $j = 1, 2$, we can rewrite (12b), (12c), (12d) as a second-order differential system in the variable x:

$$\partial_{xx}X_j - M_S X_j = 0, \quad M_S = \frac{1}{\lambda}\begin{bmatrix} d_S & -\frac{1}{\sigma} \\ 1 & \overline{d_S} \end{bmatrix}. \tag{16}$$

The variable κ appears as a parameter, and is omitted in most formulas. The boundary conditions are

$$\partial_x X^1(x_1) + pX^1(x_1) = \underline{\widehat{\mathcal{G}}}_1, \quad -\partial_x X^2 + pX^2(x_2) = \underline{\widehat{\mathcal{G}}}_2, \tag{17}$$

and the result of the iteration is

$$\underline{\widehat{\mathcal{G}}}_1' = \partial_x X^2(x_1) + pX^2(x_1), \quad \underline{\widehat{\mathcal{G}}}_2' = -\partial_x X_1(x_2) + pX^2(x_2). \tag{18}$$

For any $\kappa \in [\![0, S-1]\!]$, the matrix $M_S(\kappa)$ has two distinct eigenvalues, complex conjugate, $\mu_S(\kappa)$ and $\overline{\mu_S(\kappa)}$. It is thus diagonalizable with the eigenmatrix P into $M_S = PDP^{-1}$, with

$$D = \begin{bmatrix} \mu_S & 0 \\ 0 & \overline{\mu_S} \end{bmatrix}.$$

Define $\mathcal{X}_j = P^{-1}X_j$. Then, the iteration (16), (17), and (18), diagonalizes into

$$\partial_{xx}\mathcal{X}_j - D\mathcal{X}_j = 0, \tag{19a}$$

$$\left\{\begin{array}{l} \partial_x\mathcal{X}_1(x_1) + p\mathcal{X}_1(x_1) = P^{-1}\underline{\widehat{\mathcal{G}}}_1 := \underline{\widehat{\mathcal{H}}}_1, \\ -\partial_x\mathcal{X}^2(x_2) + p\mathcal{X}^2(x_2) = P^{-1}\underline{\widehat{\mathcal{G}}}_2 := \underline{\widehat{\mathcal{H}}}_2, \end{array}\right. \tag{19b}$$

$$\left\{\begin{array}{l} \underline{\widehat{\mathcal{H}}}_1' := P^{-1}\underline{\widehat{\mathcal{G}}}_1' = \partial_x\mathcal{X}^2(x_1) + p\mathcal{X}^2(x_1), \\ \underline{\widehat{\mathcal{H}}}_2' := P^{-1}\underline{\widehat{\mathcal{G}}}_2' = -\partial_x\mathcal{X}_1(x_2) + p\mathcal{X}^2(x_2). \end{array}\right. \tag{19c}$$

Let $z_S(\kappa)$ be the unique square root of $\mu_S(\kappa)$ with positive real part, then

$$\sqrt{D}=\begin{bmatrix} z_S(\kappa) & 0\\ 0 & \overline{z_S(\kappa)}\end{bmatrix} \text{ and } e^{\sqrt{D}x}=\begin{bmatrix} e^{z_S(\kappa)x} & 0\\ 0 & e^{\overline{z_S(\kappa)x}}\end{bmatrix}.$$

It is first easy to solve (19a) and get

$$\mathcal{X}^1=e^{\sqrt{D}x}a^1+e^{-\sqrt{D}x}b^1,\quad \mathcal{X}^2=e^{-\sqrt{D}x}a^2+e^{\sqrt{D}x}b^2.$$

Now, $\mathcal{X}^1$ and $\mathcal{X}^2$ have to vanish for $x\to-\infty$ and $x\to+\infty$, respectively (in order to be a temperate distribution), therefore $b^1=b^2=0$ and thus

$$\mathcal{X}^1=e^{\sqrt{D}x}a^1,\quad \mathcal{X}^2=e^{-\sqrt{D}x}a^2. \tag{20}$$

Inserting these expressions in the boundary iteration (19b), (19c) yields

$$\begin{aligned}&(\sqrt{D}+pI)e^{\sqrt{D}x_1}a^1=\underline{\widehat{\mathcal{H}}}_1,\quad (\sqrt{D}+pI)e^{-\sqrt{D}x_2}a^2=\underline{\widehat{\mathcal{H}}}_2,\\ &\underline{\widehat{\mathcal{H}}}_1'=(-\sqrt{D}+pI)^{-\sqrt{D}x_1}a^2,\quad \underline{\widehat{\mathcal{H}}}_2'=(-\sqrt{D}+pI)^{\sqrt{D}x_2}a^1.\end{aligned}$$

Then, the relation (15) follows by recalling that $\underline{\widehat{\mathcal{G}}}_j=P\underline{\widehat{\mathcal{H}}}_j$. □

We can now state the main result of this section.

Theorem 7 *(Semidiscrete L^2-error bounds and convergence of the OSWRM). Let $\lambda>0$ and $d>0$. There is a constant $C>0$ such that, for any initial guess $\underline{G}^0\in R^4_\#$, and for any $p>0$ and $\sigma>0$, the sequence $\underline{G}^n$ defined by (9), (11) satisfies (in error form)*

$$\|\underline{\mathcal{G}}^n\|\le C\sup_{\kappa\in[\![0,S-1]\!]}|\rho_S(\kappa,p,L)|^n\|\underline{\mathcal{G}}^0\|.$$

Furthermore, $\sup_{\kappa\in[\![0,S-1]\!]}|\rho_S(\kappa,p,L)|<1$, *therefore the sequence is convergent.*

Proof. The diagonal matrix $(\sqrt{D}+pI)$ is invertible because $\mathrm{Re}\left(z_S(\kappa)\right)$ and p are positive. Consider the matrix $G_S=(\sqrt{D}+pI)^{-1}(-\sqrt{D}+pI)e^{-\sqrt{D}L}$, given in (14). We can rewrite (15) as

$$\begin{pmatrix}\underline{\widehat{\mathcal{H}}}_1'\\ \underline{\widehat{\mathcal{H}}}_2'\end{pmatrix}=\begin{pmatrix}0 & G_S\\ G_s & 0\end{pmatrix}\begin{pmatrix}\underline{\widehat{\mathcal{H}}}_1\\ \underline{\widehat{\mathcal{H}}}_2\end{pmatrix}.$$

The result of $2n$ iterations is $(\underline{\mathcal{G}}_1^{2n},\underline{\mathcal{G}}_2^{2n})$, simply given by

$$\underline{\widehat{\mathcal{G}}}_j^{2n}=P^{-1}G_S^{2n}P\underline{\widehat{\mathcal{G}}}_j,\quad j=1,2.$$

Define $\alpha:=\sqrt{\left(\mathrm{Im}(d_S)\right)^2+\frac{1}{\sigma}}+\mathrm{Im}(d_S)$ and $\beta:=\sqrt{\left(\mathrm{Im}(d_S)\right)^2+\frac{1}{\sigma}}-\mathrm{Im}(d_S)$. Since $\mathrm{Im}(d_S)>0$, these two functions of κ are positive. Expanding the identity above gives

$$\begin{aligned}\widehat{\mathcal{G}}_j^{2n}&=\frac{\alpha\rho_S^{2n}+\beta\overline{\rho_S}^{2n}}{\alpha+\beta}\widehat{\mathcal{G}}_j^0+\frac{\beta(-\rho_S^{2n}+\overline{\rho_S}^{2n})}{\alpha+\beta}\hat{\hat{\mathcal{G}}}_j^0,\\ \hat{\hat{\mathcal{G}}}_j^{2n}&=\frac{\alpha(-\rho_S^{2n}+\overline{\rho_S}^{2n})}{\alpha+\beta}\widehat{\mathcal{G}}_j^0+\frac{\alpha\overline{\rho_S}^{2n}+\beta\rho_S{}^{2n}}{\alpha+\beta}\hat{\hat{\mathcal{G}}}_j^0.\end{aligned}$$

From this it is easy to estimate

$$|\widehat{\mathcal{G}}_j^{2n}|\le|\rho_S|^{2n}(|\widehat{\mathcal{G}}_j^0|+2|\hat{\hat{\mathcal{G}}}_j^0|),\quad |\hat{\hat{\mathcal{G}}}_j^{2n}|\le|\rho_S|^{2n}(2|\widehat{\mathcal{G}}_j^0|+|\hat{\hat{\mathcal{G}}}_j^0|).$$

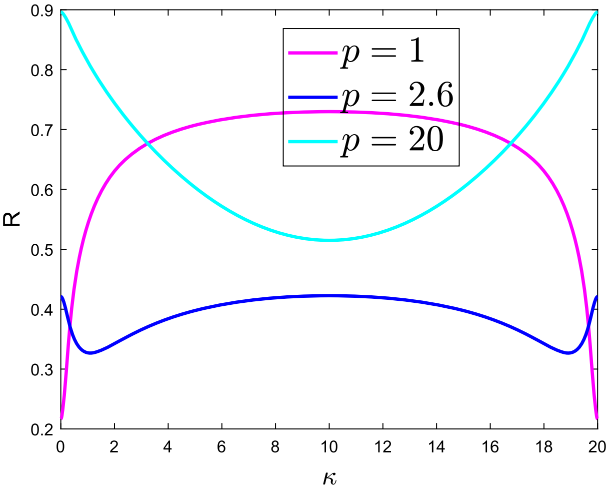
By Parseval's identity, we can conclude that

$$\|\underline{\mathcal{G}}^{2n}\|\le C\sup_{\kappa\in[\![0,S-1]\!]}|\rho_S(\kappa,p,L)|^{2n}\max(\|\mathcal{G}_j^0\|,\|\hat{\hat{\mathcal{G}}}_j^0\|).$$

For odd iterations, the error in domain j must be estimated by the previous error in domain $i\neq j$, and the result is similar.

The convergence factor $|\rho_S(\kappa,p,L)|$ is strictly smaller than 1 and the sup is taken on a compact set, therefore it is smaller than 1. □

Remark 1. The computation and the convergence proof presented in this section extend to the continuous case, using Fourier series $y(t)=\sum_{k\in\mathbb{Z}}\hat{y}_k e^{\frac{2ik\pi}{T}t}$. The relevant quantities in the notations are replaced by



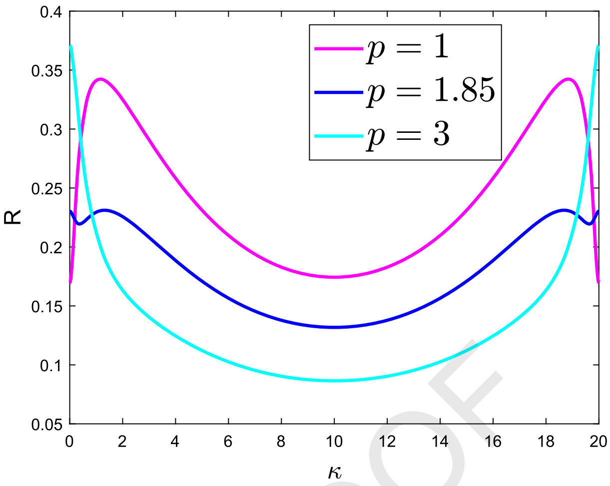


Fig. 1. Plots of R_S as a function of $\kappa \in [0, S]$ for three values of p. Left $L = 0$, right $L = \sqrt{\Delta t} = \sqrt{1/20}$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$\begin{aligned}
&d_\infty(k) = \frac{2ik\pi}{T} + d,\\
&\mu_\infty(k) = \frac{1}{\lambda}\left(\mathrm{Re}\left(d_\infty\right) + i\sqrt{\frac{1}{\sigma} + \left(\mathrm{Im}(d_\infty)\right)^2}\right) = \frac{1}{\lambda}\left(d + i\sqrt{\frac{1}{\sigma} + \left(\frac{2k\pi}{T}\right)^2}\right),\\
&z_\infty = \sqrt{\mu_\infty}, \quad \rho_\infty(k, p, L) = \frac{z_\infty(k) - p}{z_\infty(k) + p} e^{-z_\infty(k)L}.
\end{aligned} \tag{21}$$

The well-posedness and convergence analysis above concerns the overlapping case, for which no other convergence proof is available. Only slight modifications would be needed to obtain an analysis in the nonoverlapping case, see [27]. In particular, in the semidiscrete case, the convergence of nonoverlapping methods is guaranteed by the compactness of the set of possible Fourier frequencies. Thus, one can prove that the contraction factor is smaller than a constant lower than one. This is not possible in the continuous case, for which the set of Fourier frequencies is unbounded. In this case, Parseval's identity together with the dominated convergence (Lebesgue) theorem need to be used. However, the optimization study below concerns both cases.

4. Optimization of the semi-discrete Robin parameter p

By Theorem 7, the convergence speed of the algorithm is measured by the maximum over all discrete frequencies κ of the convergence factor

$$R_S(z_S(\kappa), p, L) = |\rho_S(z_S(\kappa), p, L)|^2, \text{ where } \rho_S(z_S(\kappa), p, L) = \frac{z_S(\kappa) - p}{z_S(\kappa) + p} e^{-Lz_S(\kappa)}, \tag{22}$$

with the definitions in Lemma 6. The value depends on the positive parameter p. It is always smaller than 1, but the behavior of R_S as a function of κ, and hence its maximum, depends very heavily on p, see Fig. 1, with coefficients σ, λ and d equal to 1, $T = 1$ and $S = 20$, in the nonoverlapping case $L = 0$ on the left, and in the overlapping case $L = \sqrt{\Delta t} = \sqrt{\frac{T}{S}} = \sqrt{1/20}$ on the right. We see on this example that the convergence factor of the method can be divided by 2 by choosing carefully the parameter p.

The optimal parameter p should minimize the maximum of R_S over all discrete frequencies in the range, leading to the minmax problem of finding $(p_L^*, \delta_L^*) \in \mathbb{R}_+ \times \mathbb{R}_+$ such that

$$\delta_L^* = \sup_{\kappa \in [\![0, S-1]\!]} R_S(z_S(\kappa), p_L^*, L) = \inf_{p \in \mathbb{R}_+} \sup_{\kappa \in [\![0, S-1]\!]} R_S(z_S(\kappa), p, L).$$

It is easier to extend the range in κ to the segment $[0, S-1]$, and it is the problem we study in this section. A very precise analysis in [43] shows that it makes very little difference. We will prove well-posedness (existence and uniqueness), give a precise characterization of p_L^*, using the derivative of R_S in the κ variable, and provide useful asymptotic formulas as $\Delta t \to 0$. Before stating our main results, we introduce some notations. Since $d_S(S - \kappa) = \overline{d_S(\kappa)}$, see Fig. 1, the interval in the definition of the minmax problem can be reduced to the interval $[0, \lfloor S/2 \rfloor]$. Define also $Q := \{z \in \mathbb{C}, \mathrm{Re}(z) > 0 \text{ and } \mathrm{Im}(z) > 0\}$ and $C := z_S([0, \lfloor S/2 \rfloor]) \subset Q$. Then the minmax problem to study can be written in the two equivalent forms (see (14))

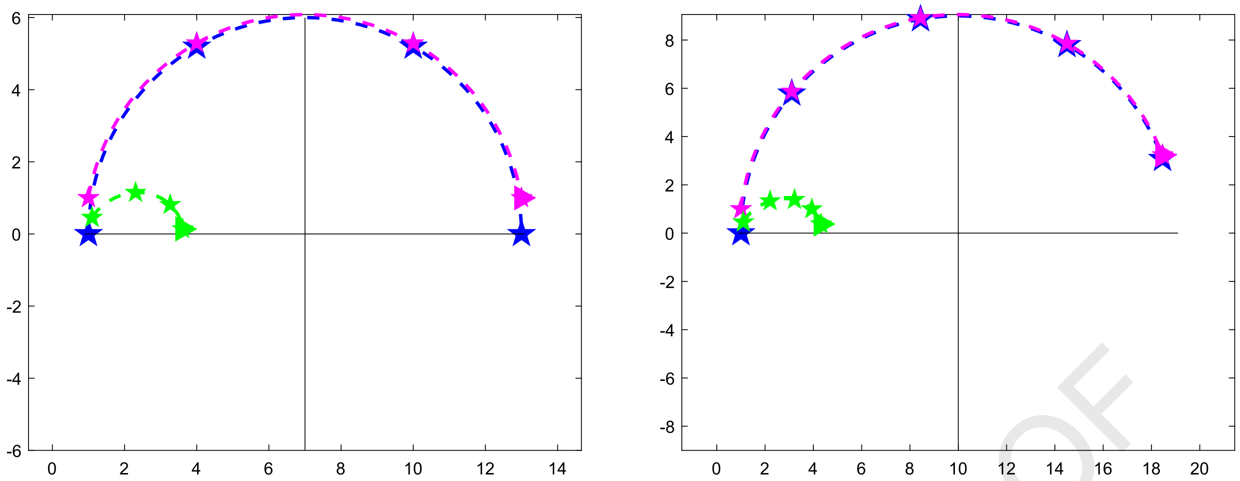


Fig. 2. Plots of d_S (blue), μ_S (magenta), and z_S (green) over $[0, \lfloor S/2 \rfloor]$ for $S = 6$ (left) and $S = 9$ (right). The arrow is the endpoint at $\lfloor \frac{S}{2} \rfloor$.

$$
\begin{aligned}
\delta_L^* &= \sup_{\kappa \in [0, \lfloor S/2 \rfloor]} R_S(z_S(\kappa), p_L^*, L) = \inf_{p \in \mathbb{R}_+} \sup_{\kappa \in [0, \lfloor S/2 \rfloor]} R_S(z_S(\kappa), p, L), \\
\delta_L^* &= \sup_{z \in C} R_S(z, p_L^*, L) = \inf_{p \in \mathbb{R}_+} \sup_{z \in C} R_S(z, p, L).
\end{aligned}
\tag{23}
$$

The image of $[0, \lfloor S/2 \rfloor]$ by the applications d_S, μ_S and z_S are plotted in Fig. 2. C is in green.

Notation 1 *(Main variables used in the proofs).*

$$
\begin{aligned}
& d_m := d_S(0) = d, \\
& d_M := d_S(\lfloor \frac{S}{2} \rfloor) = \begin{cases} d + \frac{2}{\Delta t} & \text{if } S \text{ is even}, \\ d + \frac{1}{\Delta t}(1 + e^{\frac{i\pi}{S}}) & \text{if } S \text{ is odd}. \end{cases} \\
& \mu_m := \mu_S(0) = \frac{1}{\lambda}\left(d + \frac{i}{\sqrt{\sigma}}\right), \\
& \mu_M := \mu_S(\lfloor S/2 \rfloor) = \begin{cases} \frac{1}{\lambda}\left(d + \frac{2}{\Delta t} + \frac{i}{\sqrt{\sigma}}\right) & \text{if } S \text{ is even}, \\ \frac{1}{\lambda}\left(d + \frac{1}{\Delta t}(1 + \cos\frac{\pi}{S}) + i\sqrt{\frac{1}{\sigma} + \frac{1}{\Delta t^2}\sin^2\frac{\pi}{S}}\right) & \text{if } S \text{ is odd}, \end{cases} \\
& z_m = \sqrt{\mu_m}, \ z_M = \sqrt{\mu_M}, \\
& a = \frac{1}{\lambda}\left(\frac{1}{\Delta t} + d\right), \ b = \frac{1}{\lambda}\sqrt{\frac{1}{\Delta t^2} + \frac{1}{\sigma}}.
\end{aligned}
\tag{24}
$$

Notice that in what follows, when there is no risk of confusion, the subscript S is dropped for the sake of simplicity and without loss of generality.

The main results, namely Theorem 8 and Theorem 9, are given below. Their proofs are outlined at the end of this paragraph, and they are given in the next subsections 4.1 and 4.2.

Theorem 8 *(Existence and formula for $L = 0$). Let $L = 0$. The inf-sup problem (23) has a unique solution, the best parameter and the best value are given by*

$$
p_0^* = \sqrt{\frac{\operatorname{Re}(z_m)\,|z_M|^2 - \operatorname{Re}(z_M)\,|z_m|^2}{\operatorname{Re}(z_M) - \operatorname{Re}(z_m)}}, \quad \delta_0^* = R_0(z_m, p_0^*) = R_0(z_M, p_0^*) < 1.
\tag{25}
$$

For small Δt, their asymptotics are given by

$$
p_0^* \sim \sqrt{\sqrt{\frac{2}{\lambda}}\operatorname{Re}(z_m)}\,\Delta t^{-\frac{1}{4}}, \quad \delta_0^* \sim 1 - 4\sqrt{\sqrt{\frac{\lambda}{2}}\operatorname{Re}(z_m)}\,\Delta t^{\frac{1}{4}},
\tag{26}
$$

z_m and z_M are defined in (24).

In general, for a discrete algorithm it is important to study its behavior with respect to the discretization parameter. For this reason, we give asymptotic (in Δt) results in Theorem 8 and Theorem 9, to characterize convergence of the OSWRM algorithm for Δt small. See also the discussion after Theorem 9.

In the overlapping case, the size L of the overlap is in general a few grid points. Furthermore, time and space meshes have to be chosen taking into account stability and accuracy. For the implicit scheme in consideration here, there is no stability condition, therefore the space and time mesh can be of the same magnitude. However, for accuracy, one needs rather Δt to be of the magnitude of Δx^2. For a thorough discussion on this point see [27]. Therefore, we will consider these two cases. We will use the shorthand $a \approx b$ to say that a and b are of same order as a parameter tends to 0, without specifying any constant.

Theorem 9 *(Existence and formula for $L > 0$). Let $L > 0$. The inf-sup problem (23) has a unique solution (p_L^*, δ_L^*), and $\delta_L^* < 1$. There exists L_0 such that, for $L \leq L_0$, the parameter p_L^* and the convergence factor δ_L^* have the following behavior:*

1. *If $\Delta t \approx L$,*
$$p_L^* \sim \sqrt{\sqrt{\frac{2}{\lambda}}\operatorname{Re}(z_m)}\,\Delta t^{-\frac{1}{4}}, \quad \delta_L^* \sim 1 - 4\sqrt{\sqrt{\frac{\lambda}{2}}\operatorname{Re}(z_m)}\,\Delta t^{\frac{1}{4}},$$
and the convergence factor equioscillates at endpoints z_m and z_M of C.
2. *If $\Delta t \approx L^2$,*
$$p_L^* \sim (\operatorname{Re}(z_m))^{\frac{2}{3}} L^{-\frac{1}{3}}, \quad \delta_L^* \sim 1 - 4(\operatorname{Re}(z_m))^{\frac{1}{3}} L^{\frac{1}{3}},$$
and the convergence factor equioscillates at points z_m and $z_2 \sim \sqrt{\frac{2p_L^*}{L}} e^{\frac{i\pi}{4}}$ of C.

Note that in the first case (i.e., $\Delta t \approx L$), the parameter is asymptotically equal to p_0^*: an overlap proportional to Δt does not affect the minimization problem. Note also that the overlapping algorithm with $\Delta t \approx \Delta x^2$ improves the convergence speed for small mesh, since $1 - \delta_L^*$ behaves like $\Delta t^{\frac{1}{6}}$ instead of $\Delta t^{\frac{1}{4}}$.

Before turning to the proof of the theorems, some general remarks on the geometric objects used here are in order.

According to (14), $d_S(\kappa) - (\frac{1}{\Delta t} + d) = \frac{1}{\Delta t} e^{-\frac{2i\pi}{S}}$. Therefore $d_S([0, \lfloor S/2 \rfloor])$ is an arc of the circle of center $\frac{1}{\Delta t} + d$ and radius $\frac{1}{\Delta t}$. Furthermore, $\left|\mu(\kappa) - \frac{1}{\lambda}(\frac{1}{\Delta t} + d)\right| = \frac{1}{\lambda}\sqrt{\frac{1}{\Delta t^2} + \frac{1}{\sigma}}$. Therefore, $\mu([0, \lfloor S/2 \rfloor])$ is an arc of the circle of center $\frac{1}{\lambda}\left(\frac{1}{\Delta t} + d\right)$ and radius $\frac{1}{\lambda}\sqrt{\frac{1}{\Delta t^2} + \frac{1}{\sigma}}$, joining the points μ_m and μ_M. It can be described by the angle θ:

$$\mu(\theta(\kappa)) = a + be^{i\theta(\kappa)}, \; a = \frac{1}{\lambda}\left(\frac{1}{\Delta t} + d\right), \; b = \frac{1}{\lambda}\sqrt{\frac{1}{\Delta t^2} + \frac{1}{\sigma}} \text{ are given in (24).} \tag{27}$$

Remark 2 *(Geometric properties and Cassini ovals).*

1. When κ increases from 0 to $\lfloor S/2 \rfloor$, θ decreases from θ_m to θ_M, $|\mu|$ increases from $|\mu_m|$ to $|\mu_M|$, and $|z|$ increases from $|z_m| = \sqrt{|\mu_m|}$ to $|z_M| = \sqrt{|\mu_M|}$.
2. For $z \in Q$, and positive p, $|z - p| \leq |z + p|$. Therefore, the solution p^* of the inf-sup problem for p in $\mathbb{R}$ or in $\mathbb{R}_+$ are the same, and $\delta_L^* \leq 1$.
3. For a geometric interpretation of z, note that $|\mu - a| = b$, can be rewritten as $|z^2 - a| = b$ or equivalently $|(z - \sqrt{a})(z + \sqrt{a})| = b$. Defining the foci $F_1 = -\sqrt{a}$ and $F_2 = \sqrt{a}$, we see that the product of the distances of z to F_1 and F_2 is a constant equal to b (whereas for an ellipse the sum of the distances is constant). The curves defined by this property are called Cassini ovals (Giovanni Domenico Cassini, 1680).[1] Then C is the part of the Cassini oval in Q, between $z_m = \sqrt{\mu_m}$ and $z_M = \sqrt{\mu_M}$. The Cassini ovals are quartic, thus our inf-sup problem differs from the ones in [19,27,28,31], and the arguments for the formulas are very different.

Remark 3. In the continuous case. z_∞ belongs to the hyperbola $x^2 - y^2 = d$, and the computations are slight modifications of those in [27, Theorems 5.14 and 5.18]. The asymptotic behavior is the same, with slightly different coefficients. For example in the nonoverlapping case, the optimal parameter in the PDE case is equal to

$$p_{0,\infty}^* \sim \sqrt{2\sqrt{\frac{\pi}{\lambda}}\operatorname{Re}(z_m)}\,\Delta t^{-\frac{1}{4}}, \quad \delta_0^* \sim 1 - 4\sqrt{\sqrt{\frac{\lambda}{4\pi}}\operatorname{Re}(z_m)}\,\Delta t^{\frac{1}{4}},$$

Similar differences between PDE and discretized PDE have been highlighted for elliptic or parabolic equations, see [44,45].

In Fig. 3 we compare the behavior of the convergence factor for the optimal discrete parameter and the optimal continuous parameters. The coefficients are the same as in the previous example, $\sigma = \lambda = d = 1$, $T = 1$ and $S = 20$.

[1] See https://mathcurve.com/courbes2d/cassini/cassini.shtml.

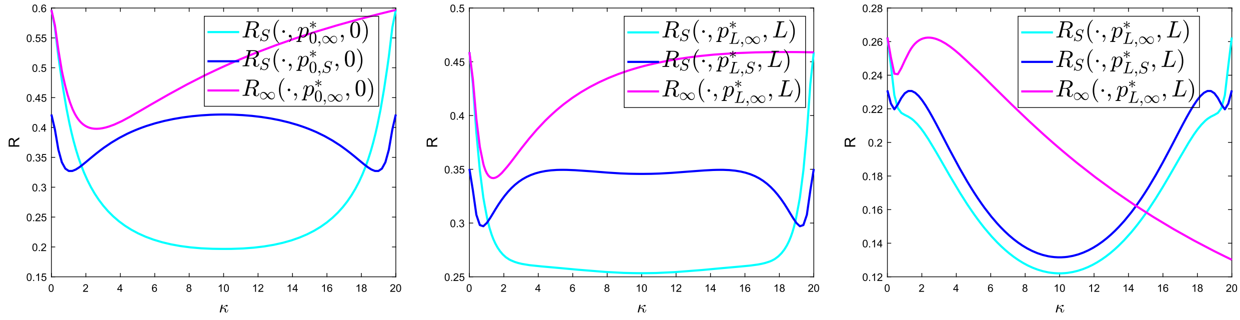


Fig. 3. Plots of the discrete convergence factor for the discrete optimized p and the continuous optimized p. Left: nonoverlaping case. Center: overlapping Case 1 ($L \approx \Delta t = 1/20$). Right: overlapping Case 2 ($L \approx \sqrt{\Delta t} = \sqrt{1/20}$).

The plot in magenta is the best convergence factor for the PDE, based on the optimization of ρ_∞ over the interval $[\![0, S]\!]$. The plot in blue is the best convergence factor for the semi-discrete equation, based on the optimization of ρ_S over the interval $[\![0, \lfloor S/2 \rfloor]\!]$. In cyan is the convergence factor of the semi-discrete equation, when using the best parameter for the PDE, the one we would have used if we hadn't performed the analysis in the paper. We clearly see the gain in using the new analysis, which can be 40% in the nonoverlapping case, to 13% in the last case, which is altogether much better than the nonoverlapping case.

Outline of the proofs

1. Prove that there are two equioscillation points z_j, that are such that $R(z_1, p^*_L, L) = R(z_2, p^*_L, L) = \delta^*_L$.
2. Identify the extremum points on C and characterize the solutions.

The first step is performed in Section 4.1. It is an easy extension of earlier results, see [19,31]. The second step requires new computations, due to the quadratic form of the curve C. It is performed in Section 4.2 for $L = 0$ and in Section 4.3 for $L > 0$.

4.1. Existence, uniqueness and equioscillation

Define the functions

$$h_L(p) = \max_{z \in C} R(z, p, L), \quad \tilde{z}(p) = h_L(p).$$

Since ρ is a continuous function on $C \times \mathbb{R}_+ \times \mathbb{R}_+$ and the maximum is taken on a compact set, h_L is well defined and continuous on $\mathbb{R}_+$. For any p the maximum is attained at some $\tilde{z}(p) \in C$.

Theorem 10. *For any $L \geq 0$, the inf-sup problem (23) has a unique solution $(p^*_L, \delta^*_L) \in \mathbb{R}_+ \times \mathbb{R}_+$, and $p^*_L \in [|z_m|, |z_M|]$. There are at least two equioscillation points z_j, that is such that*

$$R(z_1, p^*_L, L) = R(z_2, p^*_L, L) = \delta^*_L.$$

Furthermore, any strict local minimum of h_L is a global minimum.

Proof. First, we prove that for any $p < |z_m|$, $h_L(p) \geq h_L(|z_m|)$. For any $z \in C$, p_1 and p_2 in $\mathbb{R}_+$, we compute

$$R_0(z, p_1) - R_0(z, p_2) = 4\mathrm{Re}\,(z)\, \frac{(p_2 - p_1)(|z|^2 - p_1 p_2)}{|z + p_1|^2 |z + p_2|^2}. \tag{28}$$

Now, we apply this identity to $p_2 = |z_m|$, $p_1 < p_2$ and $z = \tilde{z}(p_2)$. Then

$$|\tilde{z}(p_2)|^2 - p_1 p_2 \geq |\tilde{z}(p_2)|^2 - |z_m|^2 \geq 0,$$

which together with the fact that $\mathrm{Re}\left(\tilde{z}(p_2)\right) > 0$, implies that

$$R_0(\tilde{z}(p_2), p_1) - R_0(\tilde{z}(p_2), p_2) \geq 0. \tag{29}$$

Use this for a lower bound on $h_L(p_1) - h_L(p_2)$

$$\begin{aligned} h_L(p_1) - h_L(p_2) &= \max_{z \in C} R(z, p_1, L) - R(\tilde{z}(p_2), p_2, L) \\ &\geq R(\tilde{z}(p_2), p_1, L) - R(\tilde{z}(p_2), p_2, L) \\ &= (R_0(\tilde{z}(p_2), p_1) - R_0(\tilde{z}(p_2), p_2)) e^{-2L\mathrm{Re}(\tilde{z}(p_2))} \\ &\geq 0 \text{ by (29)}. \end{aligned}$$

Therefore, we have proved that for any $p \le |z_m|$, $h_L(p) \ge h_L(|z_m|)$. An analogous computation shows that $h_L(p) \ge h_L(|z_m|)$ for $p \ge |z_m|$. This proves by compactness that the continuous function h_L has a minimum on $\mathbb{R}_+$, which is attained in the interval $[|z_m|, |z_m|]$. Finally, equioscillation, uniqueness, and the fact that strict local minimizers are global minimizers are proved exactly as in [19, 31]. □

4.2. Optimal solution in the nonoverlapping case $L = 0$, proof of Theorem 8

4.2.1. Identification of the extremum points on C

We begin by computing the derivative of R_0 with respect to θ. This is done in the following lemma.

Lemma 11 *(Derivative of the convergence factor R_0). It holds that*

$$\begin{aligned} \frac{\partial \mathrm{Re}\,(z)}{\partial\theta} &= \frac{1}{2}\frac{-(|\mu|+a)\mathrm{Im}(z)}{|z|^2}, \\ \frac{\partial R_0}{\partial\theta} &= 2p\frac{\phi(|\mu|)\mathrm{Im}(z)}{|z|^2|z+p|^4}, \text{ with } \phi(|\mu|) := b^2 + a(p^2-a) + (p^2-a)|\mu|. \end{aligned} \tag{30}$$

Proof. Recalling that $z = \sqrt{\mu}$ and noticing that $\mu'(\theta) = i(\mu - a)$, we get

$$\frac{dz}{d\theta} = \frac{1}{2\mu}\frac{dz}{d\theta} = \frac{i(\mu-a)\bar{z}}{2|z|^2} = \frac{i(|\mu|z - a\bar{z})}{2|z|^2},$$

which leads to the first formula. Compute now the derivative of ρ_0 in θ:

$$\begin{aligned} &\frac{\partial\rho_0}{\partial z} = \frac{2p}{(z+p)^2}, \quad \frac{\partial\rho_0}{\partial\theta} = \frac{\partial\rho_0}{\partial z}\frac{dz}{d\theta} = i\frac{p}{(z+p)^2}\frac{(\mu-a)\bar{z}}{2|z|^2}, \\ &\frac{\partial R_0}{\partial\theta} = 2\mathrm{Re}\left(\overline{\rho_0}\frac{\partial\rho_0}{\partial\theta}\right) = 2\mathrm{Re}\left(i\frac{\bar{z}-p}{\bar{z}+p}\frac{p}{(z+p)^2}\frac{(\mu-a)\bar{z}}{|z|^2}\right) = -2p\frac{\mathrm{Im}((\bar{\mu}-p^2)(\mu-a)\bar{z})}{|z|^2|z+p|^4}. \end{aligned}$$

Consider now the numerator of this expression. From $\mu = a + be^{i\theta}$ we find

$$(\mu-a)(\bar{\mu}-a) = b^2 \implies \bar{\mu}(\mu-a) = b^2 + a(\mu-a).$$

Therefore, recalling that $\mu = z^2$, we have

$$(\bar{\mu}-p^2)(\mu-a)\bar{z} = (b^2 - a(a-p^2) + (a-p^2)z^2)\bar{z} = (b^2 - a(a-p^2))\bar{z} + (a-p^2)|\mu|z,$$

and hence

$$\mathrm{Im}((\bar{\mu}-p^2)(\mu-a)\bar{z}) = \left(-(b^2-a(a-p^2)) + (a-p^2)|\mu|\right)\mathrm{Im}(z),$$

which leads to the claimed derivative in θ. □

Since ϕ is an affine function, it changes sign at most once, and R_0 has at most one local extremum point. For positive p, R_0 is smaller than 1 in Q. Furthermore $R_0(0) = 1$. Therefore the extremum point is a minimum. Whether it belongs to C or not, the maximum is attained at either endpoints of C, namely z_m or z_M.

4.2.2. Conclusion

By Theorem 10, the optimal solution p_0^* must produce equioscillation in at least two points on C. These points have therefore to be z_m and z_M. By expanding the equality $R_0(z_m, p) = R_0(z_M, p)$ we get

$$\frac{(\mathrm{Re}\,(z_m)-p)^2 + \mathrm{Im}(z_m)^2}{(\mathrm{Re}\,(z_m)+p)^2 + \mathrm{Im}(z_m)^2} = \frac{(\mathrm{Re}\,(z_M)-p)^2 + \mathrm{Im}(z_M)^2}{(\mathrm{Re}\,(z_M)+p)^2 + \mathrm{Im}(z_M)^2},$$

which leads to the unique positive value $\hat{p} = \sqrt{\frac{\mathrm{Re}(z_m)|z_M|^2 - \mathrm{Re}(z_M)|z_M|^2}{\mathrm{Re}(z_M)-\mathrm{Re}(z_m)}}$. Therefore $\hat{p} = p_0^*$ and the proof is complete.

4.2.3. Asymptotics in Δt

For small Δt and large S, with $S\Delta t = T$, $\mu_m = \mathcal{O}(1)$ and $\mu_M = \frac{2}{\lambda\Delta t}(1+\mathcal{O}(\Delta t))$, from which we deduce that $z_m = \mathcal{O}(1)$ and $z_M = \sqrt{\frac{2}{\lambda\Delta t}}(1+\mathcal{O}(\Delta t)) \sim \sqrt{2a}$. Replacing these in (25) gives

$$p_0^* \sim \sqrt{\frac{\mathrm{Re}\,(z_m)\,(2a) - \sqrt{2a}|z_M|^2}{\sqrt{2a} - \mathrm{Re}\,(z_m)}} \sim \sqrt{\sqrt{2a}\mathrm{Re}\,(z_m)}, \quad a = \frac{1}{\lambda}\left(\frac{1}{\Delta t} + d\right).$$

Now, a direct calculation shows that for $z \in \mathbb{C}$, the Taylor expansion of $|(1-z)/(1+z)|^2$ at $z = 0$ at order 1 is

$$\left|\frac{1-z}{1+z}\right|^2 = 1 - 4\mathrm{Re}\,(z) + o(z). \tag{31}$$

Since $p_0^* \gg 1$, we can apply it to $z = \frac{z_m}{p_0^*}$, and obtain the asymptotics for $\delta_0^* = R_0(z_m, p_0^*)$:

$$\sim 1 - 4\frac{\mathrm{Re}\,(z_m)}{p_0^*} \sim 1 - 4\frac{\mathrm{Re}\,(z_m)}{\sqrt{\sqrt{2a}\mathrm{Re}\,(z_m)}} \sim 1 - 4\frac{\sqrt{\mathrm{Re}\,(z_m)}}{\sqrt[4]{2a}},$$

which is precisely the asymptotic result in Theorem 8.

4.3. Optimal solution in the overlapping case $L > 0$, proof of Theorem 9

Consider the inf-sup problem (23). Existence and uniqueness of the solution is provided by Theorem 10. We use an asymptotic analysis for small Δt and L to characterize the solution. The proof goes in four steps:

Step 1. We show that for $(L, \Delta t, p)$ in some subset $\mathcal{S}$ of $\mathbb{R}_+^3$, the derivative of R in z has three real roots and study their asymptotic behavior. This is needed to characterize the extrema of R in Step 2.

Step 2. We show that, for $(L, \Delta t, p)$ in $\mathcal{S}$ there are at most two local maximum points (including the endpoints) of $z \to R(z, p, L)$ in C.

Step 3. By Theorem 10, a necessary condition for p_L^* to be a minimum point for h_L is equioscillation in at least two points. Hence, we perform asymptotic expansions of R, and use them to find a parameter $\hat{p}$ with $(L, \Delta t, p)$ in $\mathcal{S}$ such that R equioscillates at the two local maximum points obtained in Step 2. We compute the asymptotic expansions of $\hat{p}$ and $\sup_z R(z, \hat{p}, L)$.

Step 4. Finally, we show that $\hat{p}$ is a local strict minimum point for h_L, and conclude by Theorem 10 that $p_L^* = \hat{p}$ is the unique minimum point for h.

For fixed p, the identification of the extremum points on C starts with the derivative of R in θ.

Lemma 12 *(Derivative of the convergence factor R). Consider the polynomial*

$$\Phi(m) = Am^3 + Bm^2 + Cm + D,$$
$$A = L(1 - \tfrac{p^2}{a}),\ B = aA,\ C = 2p(p^2 - a) + Lp^2(p^2 - \tfrac{a^2-b^2}{a}),\ D = aC + 2pb^2.$$

Then

$$\frac{\partial R}{\partial \theta} = \Phi(|\mu|)\frac{\mathrm{Im}(z)}{|z|^2|z+p|^4}e^{-2L\mathrm{Re}(z)}. \tag{32}$$

Furthermore

$$\frac{\partial R}{\partial p} = -4\frac{(|\mu| - p^2)\mathrm{Re}\,(z)}{|z+p|^4}e^{-2Lz}. \tag{33}$$

Proof. The derivative in θ is based on Lemma 11,

$$\begin{aligned}\frac{\partial R}{\partial \theta} &= (\frac{\partial R_0}{\partial \theta} - 2LR_0\frac{\partial(\mathrm{Re}\,(z))}{\partial \theta})e^{-2L\mathrm{Re}(z)} \\ &= \left(2p\frac{(b^2 + a(p^2 - a) + (p^2 - a)|\mu|)}{|z|^2|z+p|^4} + L\frac{(|\mu| + a)}{|z|^2}\frac{|z-p|^2}{|z+p|^2}\right)\mathrm{Im}(z)e^{-2L\mathrm{Re}(z)}.\end{aligned} \tag{34}$$

Reduce the term in the parenthesis to the same denominator,

$$\begin{aligned}\frac{\partial R}{\partial \theta} &= \frac{2p(b^2 + a(p^2 - a) + (p^2 - a)|\mu|) + L(|\mu| + a)|z^2 - p^2|^2}{|z|^2|z+p|^4}\mathrm{Im}(z)e^{-2L\mathrm{Re}(z)} \\ &= \frac{\Phi\,\mathrm{Im}(z)}{|z|^2|z+p|^4}e^{-2L\mathrm{Re}(z)}, \\ \Phi &= 2p(b^2 + a(p^2 - a) + (p^2 - a)|\mu|) + L(|\mu| + a)|z^2 - p^2|^2.\end{aligned}$$

The last term in Φ is evaluated as $|z^2 - p^2|^2 = |\mu - p^2|^2 = |\mu|^2 + p^4 - 2p^2\mathrm{Re}\,(\mu)$. Recalling the definition of μ, we expand $|\mu - a|^2 = b^2$ as

$$|\mu|^2 + a^2 - b^2 - 2a\mathrm{Re}\,(\mu) = 0,$$

from which we extract $\mathrm{Re}\,(\mu)$ and insert into $|\mu - p^2|^2$:

$$|\mu - a|^2 = b^2 \implies |\mu - p^2|^2 = (1 - \frac{p^2}{a})|\mu|^2 + p^4 - \frac{p^2}{a}(a^2 - b^2).$$

Now, we collect the powers of $|\mu|$ to reduce Φ to a function of $|\mu|$ only:

$$\Phi(|\mu|) = 2p(b^2 + a(p^2 - a) + (p^2 - a)|\mu|) + L(|\mu| + a)((1 - \frac{p^2}{a})|\mu|^2 + p^4 - \frac{p^2}{a}(a^2 - b^2)).$$

The formula in Lemma 12 is then obtained by ordering powers of $|\mu|$. Now, we compute the derivative in p. We begin with $\frac{\partial \rho_0}{\partial p} = \frac{-2z}{(z+p)^2}$, and then write

$$\frac{\partial R_0}{\partial p} = 2\mathrm{Re}\left(\overline{\rho_0}\frac{\partial \rho_0}{\partial p}\right) = 2\mathrm{Re}\left(\frac{\bar{z}-p}{\bar{z}+p}\frac{-2z}{(z+p)^2}\right) = -4\mathrm{Re}\left(\frac{z(\bar{z}^2-p^2)}{|z+p|^4}\right) = -4\frac{(|\mu|-p^2)\mathrm{Re}(z)}{|z+p|^4},$$

which gives the claimed formula for the derivative in p. □

Remark 4 *(Sign of derivatives of R).* Since for $z \in \mathcal{C}$, $\mathrm{Im}(z) > 0$, the derivative of R in θ (Lemma 12) has the sign of $\Phi(|\mu|)$, while the derivative of R in κ has the sign of $-\Phi(|\mu|)$. Therefore, the zeros of the derivatives of R are defined by the roots of the polynomial Φ, which is a real polynomial in m with degree three. Hence, it has one or three real roots, which Lemma 13 below makes more precise.

Lemma 13 *(Roots of the polynomial Φ). Define the coefficients*

$$P = \frac{1}{A}\left(C - \frac{B^2}{3A}\right), \quad Q = \frac{1}{A}\left(D - \frac{BC}{3A} + 2\frac{B^3}{27A^2}\right), \tag{35}$$

where A, B, C, and D are defined in Lemma 12, and

$$\Delta = -(4P^3 + 27Q^2). \tag{36}$$

If $\Delta > 0$, then Φ has the three real roots

$$u = \sqrt[3]{\frac{1}{2}\left(-Q + \sqrt{\frac{-\Delta}{27}}\right)}, \quad m_k = 2\mathrm{Re}\left(e^{\frac{2ik\pi}{3}}u\right) - \frac{a}{3}, \; k = 0, 1, 2. \tag{37}$$

Here, u is any of the three cubic roots. If $\Delta < 0$, then Φ has one real root given by $m_0 = 2\mathrm{Re}(u) - \frac{a}{3}$.

Proof. We use the Cardano formula. First, we write Φ in canonical form:

$$\Phi(m) = A\left(\left(m + \frac{B}{3A}\right)^3 + P\left(m + \frac{B}{3A}\right) + Q\right),$$

where the coefficients P and Q are defined in (35). These can be rewritten, using that $B = aA$ and $D = aC + 2pb^2$, as $P = \frac{C}{A} - \frac{a^2}{3}$, $Q = \frac{2aC}{3A} + \frac{2pb^2}{A} + 2\frac{a^3}{27}$. Hence, $\Phi(m) = A\tilde{\Phi}\left(m + \frac{B}{3A}\right)$, where $\tilde{\Phi}(m) = m^3 + Pm + Q$. The discriminant of $\tilde{\Phi}$ is exactly Δ defined in (36). Hence, the result follows by the Cardano formula (see [46] and, e.g., [31]). □

We now treat separately the two cases $\Delta t \approx L$ and $\Delta t \approx L^2$.

4.3.1. Case I: $L \approx \Delta t$

We suppose for simplicity that $\Delta t = \widetilde{C}L$ for a fixed constant $\widetilde{C} > 0$. Introduce two effective parameters

$$\gamma = \frac{p^2}{a}, \quad \eta = \frac{aL}{p}. \tag{38}$$

Define the family of sets

$$\mathcal{S}(\gamma_0, \eta_0) := \{(p, L) \in (\mathbb{R}_+)^2, \gamma \le \gamma_0 \text{ and } \eta \le \eta_0\}.$$

Step 1: identification of the extremum points of $z \to R(z, p, L)$

Lemma 14 *(Roots of Φ and their asymptotics). There exists (γ_0, η_0) such that for any $(p, L) \in \mathcal{S}(\gamma_0, \eta_0)$, the third degree polynomial Φ has 3 real roots,*

$$m_1 \sim -\sqrt{\frac{ap}{2L}} \ll m_2 \sim 4p^2 \ll m_0 \sim \sqrt{\frac{ap}{2L}}. \tag{39}$$

Proof. The computations are based on the following qualitative asymptotic study. With the notations in (24),

$$\gamma \ll 1 \text{ and } \eta \ll 1 \implies \sqrt{\gamma\eta} \ll 1 \implies \sqrt{a}L \ll 1 \implies L \ll 1. \tag{40a}$$

With this knowledge,

$$L \approx \Delta t \sim \frac{1}{\lambda a}, \quad \frac{a^2 - b^2}{a} \sim \frac{2}{\lambda}. \tag{40b}$$

Furthermore,

$$Lp^2 \approx \frac{p^2}{a} = \gamma \ll 1 \quad \text{and} \quad \frac{1}{p} \sim \frac{aL}{p} = \eta \ll 1. \tag{40c}$$

To prove existence of three real roots we rely on Lemmas 12 and 13 above, and prove that there exists (γ_0, η_0) such that for any $(p, L) \in \mathcal{S}(\gamma_0, \eta_0)$, Δ is strictly positive. Using the Cardano formula, we first show that

$$P = -\frac{2ap}{L}\left(1 + \frac{\eta}{6} - \tilde{P}\right), \quad \tilde{P} \approx \eta\gamma^2, \quad Q = \frac{2a^2p}{3L}\left(1 + \frac{\eta}{9} + \tilde{Q}\right), \quad \tilde{Q} \approx \gamma. \tag{41}$$

Start with $P = \frac{C}{A} - \frac{B^2}{3A^2}$, replace A, B, C from Lemma 12:

$$P = \frac{2p(p^2 - a)}{L(1 - \frac{p^2}{a})} + \frac{Lp^2(p^2 - \frac{a^2-b^2}{a})}{L(1 - \frac{p^2}{a})} - \frac{a^2}{3} = -\frac{2ap}{L} - \frac{a^2}{3} + \frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}}.$$

By (40), the first term is equivalent to pL^{-2}, the second to L^{-2} and the third to p^4. Thus, there exists γ_0 and η_0 such that $pL^{-2} \gg L^{-2} \gg p^4$ for all $(p, L) \in \mathcal{S}(\gamma_0, \eta_0)$. Hence, we can factorize out the first term, which is the dominant term of P, and use the parameter η to write

$$P = -\frac{2ap}{L}\left(1 + \frac{\eta}{6} - \tilde{P}\right) \quad \text{with} \quad \tilde{P} = -\frac{Lp}{2a}\frac{(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}} \approx \eta\gamma^2.$$

The evaluation of Q, is a little longer. It starts similarly

$$Q = \frac{2a}{3}\left(-\frac{2ap}{L} + \frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}}\right) + \frac{2pb^2}{L(1 - \frac{p^2}{a})} + \frac{2a^3}{27}.$$

Rewrite the third term as

$$\frac{2pb^2}{L(1 - \frac{p^2}{a})} = \frac{2pb^2}{L} + \frac{2p^3b^2}{La(1 - \frac{p^2}{a})} = \frac{2pa^2}{L} + \frac{2p(b^2 - a^2)}{L} + \frac{2p^3b^2}{La(1 - \frac{p^2}{a})},$$

which yields

$$Q = \frac{2a^2p}{3L} + \frac{2a^3}{27} + \frac{2p^3b^2}{La(1 - \frac{p^2}{a})} + \frac{2a}{3}\frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}} + \frac{2p(b^2 - a^2)}{L}. \tag{42}$$

Now by (40), we can estimate all terms of the sum:

$$Q = \underbrace{\frac{2a^2p}{3L}}_{\approx pL^{-3}(1+\mathcal{O}(L))} + \underbrace{\frac{2a^3}{27}}_{\approx L^{-3}(1+\mathcal{O}(L))} + \underbrace{\frac{2p^3b^2}{La(1 - \frac{p^2}{a})}}_{\approx p^3L^{-2}(1+\mathcal{O}(L))} + \underbrace{\frac{2a}{3}\frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}}}_{\approx p^4L^{-1}(1+\mathcal{O}(L))} + \underbrace{\frac{2p(b^2 - a^2)}{L}}_{\simeq pL^{-2}},$$

and hence write, for $Q_1 \approx p^3L^{-2}$ (and recalling that $aL \approx 1$), that

$$Q = \frac{2a^2p}{3L} + \frac{2a^3}{27} + Q_1 \Rightarrow Q = \frac{2a^2p}{3L}\left(1 + \frac{\eta}{9} + Q_1\frac{3L}{2a^2p}\right), \quad \tilde{Q} := Q_1\frac{3L}{2a^2p} \approx \frac{p^2}{a^2L} \approx \gamma,$$

and (41) is now proved.

Thus, the discriminant Δ can be written as

$$\Delta = -(4P^3 + 27Q^2) = 4\left(\frac{2ap}{L}\right)^3\left(1 + \frac{\eta}{6} - \tilde{P}\right)^3 - 3\left(\frac{2a^2p}{L}\right)^2\left(1 + \frac{\eta}{9} + \tilde{Q}\right)^2.$$

Factorizing out the first coefficient, using that $\eta = \frac{aL}{p}$, yields

$$\Delta = 4\left(\frac{2ap}{L}\right)^3\left(\left(1 + \frac{\eta}{6} - \tilde{P}\right)^3 - \frac{3\eta}{8}\left(1 + \frac{\eta}{9} + \tilde{Q}\right)^2\right).$$

Expanding in η at first order gives

$$\Delta = 4\left(\frac{2ap}{L}\right)^3\left(1 + \frac{\eta}{8} + \tilde{\Delta}\right), \quad \tilde{\Delta} = o(\eta\gamma). \tag{43}$$

This proves that there exists (γ_0, η_0) such that $\Delta > 0$ for any $(p, L) \in \mathcal{S}(\gamma_0, \eta_0)$. Hence, Lemma 13 guarantees that Φ has three real roots. We now estimate the asymptotic behavior of the three roots. First compute by (43)

$$\sqrt{\frac{-\Delta}{27}} = 2i\left(\frac{2ap}{3L}\right)^{\frac{3}{2}}\left(1+\frac{\eta}{16}+o(\eta)\right),$$

from which we get from (41)

$$\frac{1}{2}\left(-Q+\sqrt{\frac{-\Delta}{27}}\right) = i\left(\frac{2ap}{3L}\right)^{\frac{3}{2}}\left(1+i\frac{\sqrt{3}}{2}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right).$$

Then, by the definition of u in (37) we obtain ($\sqrt[3]{i}=e^{i\frac{\pi}{6}}$)

$$u=\sqrt[3]{\frac{1}{2}\left(-Q+\sqrt{\frac{-\Delta}{27}}\right)} = e^{i\frac{\pi}{6}}\left(\frac{2ap}{3L}\right)^{\frac{1}{2}}\left(1+\frac{i}{2\sqrt{3}}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right).$$

This allows us to get the asymptotics for $\tilde{m}_0$:

$$\tilde{m}_0 = 2\mathrm{Re}\,(u) = \left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1-\frac{1}{6}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right).$$

Then, using $\frac{2\pi}{3}+\frac{\pi}{6}=\pi-\frac{\pi}{6}$, we can write

$$ju = e^{\frac{2i\pi}{3}}u = -e^{-i\frac{\pi}{6}}\left(\frac{2ap}{3L}\right)^{\frac{1}{2}}\left(1+\frac{i}{2\sqrt{3}}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right),$$

and we find the asymptotics for $\tilde{m}_1$:

$$\tilde{m}_1 = 2\mathrm{Re}\,(ju) = -\left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1+\frac{1}{6}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right).$$

Now, $m_j=\tilde{m}_j-\frac{B}{3A}=\tilde{m}_j-\frac{a}{3}$. Since $a=\left(\frac{ap}{L}\right)^{\frac{1}{2}}\sqrt{\eta}$, and hence $\frac{a}{3}=\frac{2}{3}\left(\frac{ap}{2L}\right)^{\frac{1}{2}}\sqrt{\frac{\eta}{2}}$, we obtain

$$m_0 = \left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1-\left(\frac{1}{6}+\frac{2}{3}\right)\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right) = \left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1-\frac{5}{6}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right),$$
$$m_1 = -\left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1+(\frac{1}{6}+\frac{2}{3})\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right) = -\left(\frac{ap}{2L}\right)^{\frac{1}{2}}\left(1+\frac{5}{6}\sqrt{\frac{\eta}{2}}+o(\sqrt{\eta})\right).$$

For m_2, we use the product of the roots, $m_0m_1m_2=-\frac{D}{A}$. The same calculation as for Q shows that $\frac{D}{A}\sim\frac{2ap^3}{L}$. This implies the asymptotic equality

$$-\frac{2ap^3}{L}\sim-\frac{ap}{2L}m_2 \implies m_2\sim 4p^2.$$

The assumption on p proves that $m_1 \ll m_2$. This concludes the proof of the lemma. □

Step 2: the local extrema of R

Lemma 15 *(Local extrema of R). There exists (γ_0,η_0) such that for any $(p,L)\in\mathcal{S}(\gamma_0,\eta_0)$,*

$$\sup_{z\in C} R(z,p,L) = max\left(R(z_m,p,L), R(z_M,p,L)\right). \tag{44}$$

Proof. For $\gamma=\frac{p^2}{a}$ small, $A=L\left(1-\frac{p^2}{a}\right)$ is positive. Therefore, by Lemma 12 and Remark 4, the derivative of R in θ is positive for $|\mu|$ large. Thus, by Remark 2, the derivative of R with respect to κ is negative for $|\mu|$ large. Hence, since the three roots are $m_1 \ll m_2 \ll m_0$ (by Lemma 14), m_1 and m_0 are maximum points, while m_2 is the only minimum point. From the definitions (27) and (24), $m_m=|\mu_m|$ and $m_M=|\mu_M|\sim 2a$. Then

$$\frac{m_M}{m_2}\sim\sqrt{\frac{2a}{p^2}\frac{2}{\gamma}}\gg 1, \quad \frac{m_M}{m_0}\sim\sqrt{\frac{aL}{p}}=2\sqrt{2\eta}\ll 1,$$

and hence $m_1 < m_m \ll m_2 \ll m_M \ll m_0$. Therefore, there is no local maximum point inside the interval $[m_m, m_M]$, and the maximum points are either of the endpoints of the interval. Thus, the maximum points of R are attained at the extrema of C. □

Remark 5. Notice that the two extremum points z_m and z_M obtained in Lemma 15 coincide with the ones of the non-overlapping case (see Section 4.2): the overlap does not intervene in the regime $L \approx \Delta t$. The situation will be different for the case $L \approx \Delta t^{\frac{1}{2}}$ studied in Section 4.3.2.

Step 3 By the general results, the best parameter must make the convergence factor equioscillate at z_m and z_M. Therefore we now want to prove that the function $\Psi_L(p) = R(z_m, p, L) - R(z_M, p, L)$ vanishes in one point in the range defined by Lemma 15.

First compute the asymptotics of the convergence factor at the endpoints z_m and z_M.

Start with z_m. Applying (31) to $z = \frac{z_m}{p}$, with p sufficiently large for (p, L) to be in $\mathcal{S}(\gamma_0, \eta_0)$, we obtain, since $p^{-1} \approx \eta$,

$$R_0(z_m, p) = 1 - 4\frac{\mathrm{Re}\left(z_m\right)}{p} + o(\eta).$$

Thus, since $e^{-2L\mathrm{Re}(z_m)} \sim 1 - 2L\mathrm{Re}\left(z_m\right)$ and $L \ll p^{-1}$, we get

$$R(z_m, p, L) = 1 - 4\frac{\mathrm{Re}\left(z_m\right)}{p} + o(\eta). \tag{45}$$

Consider now z_M. Since $z_M = \sqrt{2a}(1 + o(L))$, $p/z_M \sim \sqrt{\frac{p^2}{2a}} \sim \sqrt{\frac{\gamma}{2}} \ll 1$. Hence, applying (31) to $z = \frac{p}{z_M}$, we get

$$R_0(z_M, p) = 1 - 4\mathrm{Re}\left(\frac{p}{z_M}\right) + o(\sqrt{\gamma}) = 1 - 4\sqrt{\frac{\gamma}{2}} + o(\sqrt{\gamma}).$$

As for the exponential term, since $L\sqrt{a} \approx \sqrt{L} \ll 1$, we have $e^{-2L\mathrm{Re}(z_M)} \sim 1 - 2L\sqrt{2a}$. Comparing $\sqrt{\gamma}$ and $L\sqrt{a}$, we get $\frac{\sqrt{\gamma}}{L\sqrt{a}} \approx \sqrt{\frac{\gamma}{L}} = \sqrt{\frac{p^2}{aL}} \approx p \gg 1$. Hence, we obtain

$$R(z_M, p, L) = \left(1 - 4\frac{p}{\sqrt{2a}}\right) + o(\sqrt{\gamma}). \tag{46}$$

Using the two expansions (45) and (46), we can evaluate Ψ_L. The parameters (γ_0, η_0) are defined by Lemma 15. Thus, we have

$$\forall (p, L) \in \mathcal{S}(\gamma_0, \eta_0), \quad \Psi_L(p) = 4\left(\frac{p}{\sqrt{2a}} - \frac{\mathrm{Re}\left(z_m\right)}{p}\right) + o(\eta) + o(\sqrt{\gamma}).$$

For $(p, L) \in \mathcal{S}(\gamma_0, \eta_0)$, if $p^2 \gg \sqrt{a}$, then $\Psi_L(p) > 0$, and if $p^2 \ll \sqrt{a}$ then $\Psi_L(p) < 0$. Then, since $\mathcal{S}(\gamma_0, \eta_0)$ is convex, by the mean value theorem, there exists $(\hat{p}, L) \in \mathcal{S}$ such that $\Psi_L(\hat{p}) = 0$. It is given asymptotically by

$$\hat{p} \sim \sqrt{\mathrm{Re}\left(z_m\right)\sqrt{2a}}, \quad \delta_L^* \sim 1 - 4\frac{\mathrm{Re}\left(z_m\right)}{p_L^*}.$$

Step 4 To finish the proof, we need to show that $p_L^* = \hat{p}$.

Lemma 16. *There exists L_0 such that for any $L \le L_0$ and $\Delta t \approx L$, $p_L^* = \hat{p}$.*

Proof. Choose L_0 and $L \le L_0$ such that $(L, \hat{p}) \in \mathcal{S}(\gamma_0, \eta_0)$ for some (γ_0, η_0). Then $h_L(\hat{p}) = \max(R(z_m, \hat{p}, L), R(z_M, \hat{p}, L))$.

By Theorem 10, we only need to show that $\hat{p}$ is a strict local minimum point for h_L. This is equivalent to showing that there exists $\varepsilon > 0$ such that for $p = \hat{p} + \varepsilon\xi$ with $|\xi| \le 1$,

$$\sup_{z \in C} R(z, \hat{p} + \varepsilon\xi, L) > \sup_{z \in C} R(z, \hat{p}, L) = R(z_m, \hat{p}, L) = R(z_M, \hat{p}, L).$$

By continuity, for ε small enough, $h_L(\hat{p} + \varepsilon\xi)$ is still the maximum of the values at z_m and z_M. It is then sufficient to prove that

$$\max(R(z_m, \hat{p} + \varepsilon\xi, L), R(z_M, \hat{p} + \varepsilon\xi, L)) > R(z_m, \hat{p}, L) = R(z_M, \hat{p}, L). \tag{47}$$

By the Taylor-Lagrange formula, there exists $0 < \delta < 1$ such that for $z = z_m$ or z_M, for any $\xi \in [-1, 1] \setminus \{0\}$,

$$R(z, \hat{p} + \varepsilon\xi, L) = R(z, \hat{p}, L) + \varepsilon\xi\frac{\partial R}{\partial p}(z, \hat{p} + \delta\varepsilon\xi, L).$$

Use now the derivative of R in p computed in (33)

$$\frac{\partial R}{\partial p} = -4\frac{(|\mu| - p^2)\mathrm{Re}\,(z)}{|z + p|^4}e^{-L\mathrm{Re}(z)}.$$

Since $\mathrm{Re}\,(z) > 0$, the sign of $\frac{\partial R}{\partial p}(z, p, L)$ is that of $p^2 - |z|^2$. Since $|z_M| \ll \hat{p}$ and $|z_M| \gg \hat{p}$) in $\mathcal{S}(\gamma_0, \eta_0)$, $\frac{\partial R}{\partial p}(z, p, L)$ is strictly positive for $z = z_m$ and strictly negative for $z = z_M$, when $p = \hat{p} + \delta\varepsilon\xi$ for ε so small that the asymptotic behavior of $\hat{p}$ is preserved. Therefore, for positive ξ, $R(z_m, \hat{p}+\varepsilon\xi, L) > R(z_m, \hat{p}, L)$, and for negative ξ, $R(z_M, \hat{p}+\varepsilon\xi, L) > R(z_M, \hat{p}, L)$. This proves (47) and terminates the proof of the theorem in the first case. □

4.3.2. Case II: $L \approx \Delta t^{\frac{1}{2}}$

For simplicity, we suppose that there exists a $\widetilde{C} > 0$ such that $\Delta t = \widetilde{C} L^2$, which implies that $aL^2 \approx 1$. In this case, the two effective parameters are

$$\gamma = \frac{p^2}{a}, \quad \zeta = \frac{1}{\eta} = \frac{p}{aL}.$$

Moreover, $aL^2 \approx 1$ implies that $\zeta \approx pL$ and $\gamma \approx \zeta^2$.

<u>Steps 1 and 2</u> Let us define the family of sets

$$\mathcal{S}(\gamma_0, \zeta_0) = \{(p, L) \in (\mathbb{R}_+)^2, \gamma \le \gamma_0 \text{ and } \zeta \le \zeta_0\}.$$

Lemma 17. *There exists (γ_0, ζ_0) such that for any $(p, L) \in \mathcal{S}(\gamma_0, \zeta_0)$, the third degree polynomial Φ has 3 real roots, with the asymptotic behavior*

$$m_0 = -a(1 + 2\zeta o(\zeta)) \ll m_1 = a\gamma(1 + o(1)) \ll m_2 = 2a\zeta(1 + o(\zeta)). \tag{48}$$

In particular, one has that

$$m_0 \sim -a \ll m_m \ll m_1 \sim p^2 \ll m_2 \sim \frac{2p}{L} \ll m_M \sim 2a. \tag{49}$$

Therefore, the convergence factor has one maximum point at $z_2 \sim \sqrt{\frac{p}{L}}(1+i)$, which is the only point on C with $|z_2|^2 = m_2$, and

$$\sup_{z \in C} R(z, p, L) = max(R(z_m, p, L), R(z_2, p, L)). \tag{50}$$

Proof. Note that for small γ_0 and ζ_0, since $\zeta \approx pL$ and $\gamma \approx (pL)^2$, we have small pL and $\gamma \approx \zeta^2$. We proceed as in the previous case, starting with the asymptotic behaviors of P and Q, we prove that

$$P = -\frac{a^2}{3}(1 + 6\zeta - 3\gamma^2 + o(\gamma^2)), \quad Q = \frac{2a^3}{27}(1 + 9\zeta + 27\gamma\zeta + \mathcal{O}(\zeta^4)). \tag{51}$$

Recalling P from Lemma 13, we write

$$P = -\frac{2ap}{L} - \frac{a^2}{3} + \frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}} = -\frac{2ap}{L} - \frac{a^2}{3} + p^4 + P_1, \quad P_1 = \frac{p^2(\frac{p^2}{a} - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}}.$$

The first term is now of the magnitude of pL^{-3}, the second of L^{-4}, the third of p^4, and the last of p^2. Then, for small γ_0 and ζ_0, we have $L^{-4} \gg pL^{-3} \gg p^4 \gg p^2$ (recalling that p is large for L small). Therefore, we can factorize out a^2 in P:

$$P = -\frac{a^2}{3}\left(1 + \frac{6p}{aL} - \frac{3p^4}{a^2} - \frac{3}{a^2}P_1\right),$$

and estimate the last term as $\frac{3}{a^2}P_1 \approx \frac{p^2}{a^2} = o(\gamma^2)$, which yields

$$P = -\frac{a^2}{3}(1 + 6\zeta - 3\gamma^2 + o(\gamma^2)).$$

Consider now Q and recall its equivalent expression (42) obtained in the proof of Lemma 14: $Q = \frac{2a^2p}{3L} + \frac{2a^3}{27} + \frac{2p^3b^2}{La(1-\frac{p^2}{a})} + \frac{2a}{3}\frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1-\frac{p^2}{a}}$.

Write now

$$\frac{2p^3b^2}{La(1 - \frac{p^2}{a})} = \frac{2p^3a}{L} + \frac{2p^3}{L}\frac{\frac{p^2}{a} - \frac{a^2-b^2}{a}}{1 - \frac{p^2}{a}} \approx \frac{2p^3a}{L} + p^3L^{-1}(1 + \mathcal{O}(L)),$$

which allows us to obtain

$$Q = \underbrace{\frac{2a^2p}{3L}}_{\approx pL^{-5}(1+\mathcal{O}(L))} + \underbrace{\frac{2a^3}{27}}_{\approx L^{-6}(1+\mathcal{O}(L))} + \underbrace{\frac{2p^3a}{L}}_{\approx p^3L^{-3}(1+\mathcal{O}(L))} + \underbrace{\frac{2p^3}{L}\frac{\frac{p^2}{a} - \frac{a^2-b^2}{a}}{1 - \frac{p^2}{a}}}_{\approx p^3L^{-1}(1+\mathcal{O}(L))} + \underbrace{\frac{2a}{3}\frac{p^2(p^2 - \frac{a^2-b^2}{a})}{1 - \frac{p^2}{a}}}_{\approx p^4L^{-2}(1+\mathcal{O}(L))}.$$

Recalling that $\zeta \approx pL$, we keep the first three terms and group the others in Q_1, to obtain

$$Q = \frac{2a^3}{27} + \frac{2a^2 p}{3L} + \frac{2p^3 a}{L} + Q_1 = \frac{2a^3}{27}(1 + 9\zeta + 27\gamma\zeta + \tilde{Q})$$

with $\tilde{Q} = \frac{27}{2a^3} Q_1 = \mathcal{O}(\zeta^4)$. Therefore, recalling (36), we obtain

$$\Delta = \frac{4a^6}{27}\left((1+6\zeta-3\gamma^2+o(\gamma^2))^3-(1+9\zeta+27\gamma\zeta+\mathcal{O}(\zeta^4))^2\right)=4a^6\zeta^2\left(1+\frac{2}{\zeta}(4\zeta^2-\gamma)+o(\zeta)\right).$$

Since $\Delta > 0$, $\tilde{\Phi}$ has three real roots that we can now compute asymptotically. For this purpose, we first calculate

$$\frac{1}{2}\left(-Q + \sqrt{\frac{-\Delta}{27}}\right) \sim -\frac{a^3}{27}\left(1 + 9\zeta + 27\gamma\zeta + \mathcal{O}(\zeta^4)\right) + \frac{i}{\sqrt{27}} a^3 \zeta \left(1 + \frac{1}{\zeta}(4\zeta^2 - \gamma) + o(\zeta)\right).$$

We factorize out $-\frac{a^3}{27}$, to obtain

$$\begin{aligned}\frac{1}{2}\left(-Q + \sqrt{\frac{-\Delta}{27}}\right) &\sim -\frac{a^3}{27}\left(1 + 9\zeta + 27\gamma\zeta - i\sqrt{27}\zeta\left(1 + \frac{1}{\zeta}(4\zeta^2 - \gamma)\right) + o(\zeta)\right)\\ &\sim -\frac{a^3}{27}\left(1 + 3(3 - i\sqrt{3})\zeta - 3i\sqrt{3}(4\zeta^2 - \gamma) + o(\zeta^2)\right).\end{aligned}$$

Thus, we obtain u as the cubic root of this expression, defined by

$$u = -\frac{a}{3}(1 + (3 - i\sqrt{3})\zeta + 2(-3 + i\sqrt{3})\zeta^2 + i\sqrt{3}\gamma + o(\zeta^2)),$$

and write for easiness

$$u = -\frac{a}{3}\tilde{u}, \quad \mathrm{Re}\,(\tilde{u}) = 1 + 3\zeta - 6\zeta^2, \quad \mathrm{Im}(\tilde{u}) = \sqrt{3}(-\zeta + 2\zeta^2 + \gamma).$$

From u we compute $\tilde{m}_k$, $k, = 0, 1, 2$:

$$\begin{aligned}\tilde{m}_0 &= 2\mathrm{Re}\,(u) = -\frac{2a}{3}(1 + 3\zeta - 6\zeta^2 + o(\zeta^2)),\\ \tilde{m}_1 &= 2\mathrm{Re}\left(e^{\frac{2i\pi}{3}} u\right) = -\frac{a}{3}\tilde{u}(-1 + i\sqrt{3}) = -\frac{a}{3}(-\mathrm{Re}\,(\tilde{u}) - \sqrt{3}\mathrm{Im}(\tilde{u})) = \frac{a}{3}(1 + 3\gamma + o(\zeta^2)),\\ \tilde{m}_2 &= 2\mathrm{Re}\left(e^{\frac{42i\pi}{3}} u\right) = -\frac{a}{3}\tilde{u}(-1 - i\sqrt{3}) = -\frac{a}{3}(-\mathrm{Re}\,(\tilde{u}) + \sqrt{3}\mathrm{Im}(\tilde{u})) = \frac{a}{3}(1 + 6\zeta - 12\zeta^2 + 3\gamma + o(\zeta^2)).\end{aligned}$$

Now, $m_j = \tilde{m}_j - \frac{a}{3}$ gives $m_0 = -a(1 + 2\zeta + o(\zeta))$, $m_1 = a\gamma(1 + o(1))$, and $m_2 = 2a\zeta(1 + o(\zeta))$. The assumption on p proves that $m_1 \ll m_2$. This concludes the proof of (48). In contrast to the previous case, both m_1 (local minimum) and m_2 (local maximum) belong to the interval $[m_m, m_M]$. Notice that z_2 is now the only point on C such that $|z_2|^2 = m_2$. We recover it by

$$\begin{array}{ll} |\mu|^2 - 2a\mathrm{Re}\,(\mu) = b^2 - a^2, & \mu \text{ belongs to the circle}\\ |z| = \sqrt{|\mu|}, \quad 2\mathrm{Re}\,(z)^2 = |\mu| + \mathrm{Re}\,(\mu) & z = \sqrt{\mu}. \end{array}$$

Extract $\mathrm{Re}\,(\mu)$ from the first line and replace it into the second to obtain the asymptotics,

$$2\mathrm{Re}\left(z_2\right)^2 = m_2(1 + \zeta + o(\zeta)), z_2 \sim \sqrt{\frac{m_2}{2}}(1 + i) = \sqrt{m_2}e^{i\frac{\pi}{4}}, \quad m_2 \sim \frac{2p}{L},$$

which concludes our proof. □

Step 3 Compute now the convergence factors at the points z_2 and z_m. First, we have

$$R(z_m, p, L) = 1 - 4\frac{\mathrm{Re}\left(z_m\right)}{p} + o(p^{-1}).$$

Next, noticing that $\frac{p}{z_2} \sim \frac{p}{\sqrt{m_2}}e^{-i\frac{\pi}{4}} \implies \mathrm{Re}\left(\frac{p}{z_2}\right) \sim \frac{p}{\sqrt{2m_2}} = \frac{\sqrt{Lp}}{2} \approx \sqrt{\zeta} \ll 1$ (and recalling $\left|\frac{1-z}{1+z}\right|^2 = 1 - 4\mathrm{Re}\,(z) + o(z)$ for $|z| \ll 1$) we obtain

$$R_0(z_2, p) = 1 - 4\mathrm{Re}\left(\frac{p}{z_2}\right) + o(\sqrt{\gamma}) = 1 - 2\sqrt{Lp} + o(\sqrt{\gamma}).$$

Now, using that

$$L\mathrm{Re}\left(z_2\right) \sim L\sqrt{\frac{p}{L}} = \sqrt{Lp} \implies e^{-2L\mathrm{Re}(z_2)} \sim 1 - 2\sqrt{Lp} + o(\sqrt{\gamma}),$$

we can evaluate R at point z_2:

$$R(z_2, p, L) = R_0(z_2, p)e^{-2L\mathrm{Re}(z_2)} = 1 - 4\sqrt{Lp} + o(\zeta).$$

Introduce the function $\Psi_L(p) = R(z_m, p, L) - R(z_2, p, L)$. For $(L, p) \in \mathcal{S}(\gamma_0, \zeta_0)$,

$$\Psi_L(p) = \sqrt{Lp} - \frac{\operatorname{Re}(z_m)}{p} + o(\zeta) + o(\zeta)o(p^{-1}). \tag{52}$$

Define $\hat{p}^{as} = \operatorname{Re}(z_m)^{\frac{2}{3}} L^{-\frac{1}{3}}$ by annihilation of the first term in the expansion. Now, we notice that

- For any $(L, p) \in \mathcal{S}(\gamma_0, \zeta_0)$ with $p \gg \hat{p}^{as}$, it holds that $\Psi_L(p) > 0$.
- For any $(L, p) \in \mathcal{S}(\gamma_0, \zeta_0)$ with $p \ll \hat{p}^{as}$, it holds that $\Psi_L(p) < 0$.

Then, by the mean value theorem, there exists $\hat{p}$ such that $(L, \hat{p}) \in \mathcal{S}(\gamma_0, \zeta_0)$ such that $\Psi_L(\hat{p}) = 0$. Recalling (52), one has that it is given asymptotically by $\hat{p} \sim \hat{p}^{as}$.

Step 4 To finish the proof, we need to show that $p^*_L = \hat{p}$.

Lemma 18. *There exists L_0 such that for any $L \le L_0$ and $\Delta t \approx L$, $p^*_L = \hat{p}$.*

Proof. With the asymptotic behavior of $\hat{p}$, we can choose L_0 and $L \ge L_0$ such that $(L, \hat{p}) \in \mathcal{S}(\gamma_0, \eta_0)$ for some (γ_0, η_0). Then $h_L(\hat{p}) = \max(R(z_m, \hat{p}, L), R(z_2(\hat{p}), \hat{p}, L))$, where we have stressed the fact that z_2 depends on p and is the largest root of Φ. By Theorem 10, we only need to show that $\hat{p}$ is a strict local minimum point for h_L. This is equivalent to showing that there exists $\varepsilon > 0$ such that for $p = \hat{p} + \varepsilon\xi$ with $|\xi| \le 1$,

$$\sup_{z \in C} R(z, \hat{p} + \varepsilon\xi, L) \; > \; \sup_{z \in C} R(z, \hat{p}, L) = R(z_m, \hat{p}, L) = R(z_2(\hat{p}), \hat{p}, L).$$

By continuity, for ε small enough, h_L is still the maximum of the values at z_m and z_2. It is then sufficient to prove that

$$\max(R(z_m, \hat{p} + \varepsilon\xi, L), R(z_2(\hat{p} + \varepsilon\xi), \hat{p} + \varepsilon\xi, L)) \; > R(z_m, \hat{p}, L) = R(z_2(\hat{p}), \hat{p}, L). \tag{53}$$

By the Taylor-Lagrange formula, there exists $0 < \delta < 1$ such that for any $\xi \in [-1, 1] \setminus \{0\}$,

$$R(z_m, \hat{p} + \varepsilon\xi, L) = R(z_m, \hat{p}, L) + \varepsilon\xi \frac{\partial R}{\partial p}(z_m, \hat{p} + \delta\varepsilon\xi, L).$$

By Lemma 11, since $\operatorname{Re}(z) > 0$, the sign of $\frac{\partial R}{\partial p}(z_m, p, L)$ is that of $p^2 - |z_M|^2$. By the asymptotics above, it is strictly positive at $p = \hat{p} + \delta\varepsilon\xi$ for ε so small that the asymptotic behavior of $\hat{p}$ is preserved. Now, we write

$$R(z_2(\hat{p} + \varepsilon\xi), \hat{p} + \varepsilon\xi, L) = R(z_2(\hat{p}), \hat{p}, L) + \varepsilon\xi \frac{\partial}{\partial p} R(z(p), p, L)|_{p = \hat{p} + \delta\varepsilon\xi},$$

and

$$\frac{\partial}{\partial p} R(z_2(p), p, L) = \frac{\partial R}{\partial p}(z_2(p), p, L) + \frac{\partial z}{\partial p}(p) \frac{\partial R}{\partial z}(z_2(p), p, L).$$

By definition of $z_2(p)$, $\frac{\partial R}{\partial z}(z_2(p), p, L) = 0$. Hence, $\frac{\partial}{\partial p} R(z_2(p), p, L) = \frac{\partial R}{\partial p}(z_2(p), p, L)$. The sign of $\frac{\partial R}{\partial p}(z_2(p), p, L)$ is that of $p^2 - |z_2(p)|^2$. By the asymptotics above, it is equal to $p^2 - m_2 \ll 0$ for $p = \hat{p} + \delta\varepsilon\xi$ for ε so small that the asymptotic behavior of $\hat{p}$ is preserved. Therefore, for positive ξ, $R(z_m, \hat{p} + \varepsilon\xi, L) > R(z_m, \hat{p}, L)$, and for negative ξ, $R(z_2(\hat{p} + \varepsilon\xi), \hat{p} + \varepsilon\xi, L) > R(z_2(\hat{p}), \hat{p}, L)$. This proves (47) and terminates the proof of the theorem in the second case. □

5. Numerical experiments

In this section, we present results of numerical experiments illustrating our theoretical analysis. In particular, we consider in all our experiments a bounded domain $\Omega = (-4 + L, 4)$, where L is the overlap, homogeneous Dirichlet conditions are imposed on the boundary of Ω, and the target state is

$$y_Q(t, x) = \left[(1 + t) \sin(\pi t) \left(e^{-8(x-1-L)^2} + e^{-8(x+1)^2} - e^{-8(1-\frac{L}{2})^2} - e^{-8(3+\frac{L}{2})^2} \right) \right]^+,$$

with $[\cdot]^+ = \max\{\cdot, 0\}$. We choose the domain $\Omega = (-4 + L, 4)$ in order to have the subdomains of the same size: $\Omega_1 = (-4 + L, L)$ and $\Omega_2 = (0, 4)$. Moreover, we set $T = 1.0$, $\lambda = 0.3$, $d = 0.5$, and $\sigma = 10^{-6}$. If one solves this problem, then the results of Fig. 4 are obtained. This figure shows the optimal control function (left panel), optimal state (middle panel), and the target state (right panel).

Now, we wish to study the convergence of the OSWRM. To do so, we first run this method till the relative error in terms of Robin traces (measured at the interfaces) becomes lower than the tolerance $\tau = 10^{-10}$. We choose the spatial discretization step $\Delta x = 0.005$ and the size of the overlap $L = 2\Delta x$. Hence, Ω is discretized with $N = 7999$ points, while the time interval $[0, T]$ with $S = 101$ equidistant points. The initialization g^0_L, $\tilde{g}^0_L$, g^0_0 and $\tilde{g}^0_0$ are chosen randomly, but satisfying the periodicity conditions, while the Robin parameter is set to $p = 1.0$. The error decay is shown in Fig. 5. In particular, we observe the decay of the L^2-norm of the

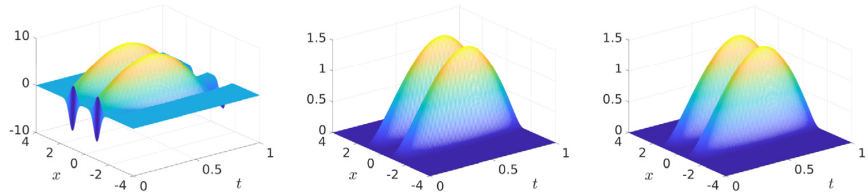


Fig. 4. Optimal control (left) and state (middle) for $\sigma = 10^{-6}$ and target (right).

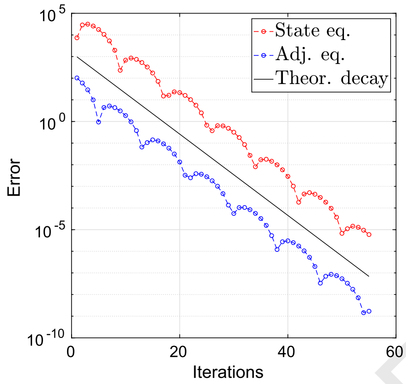


Fig. 5. Error decay of the OSWRM. ($\tau = 10^{-10}$).

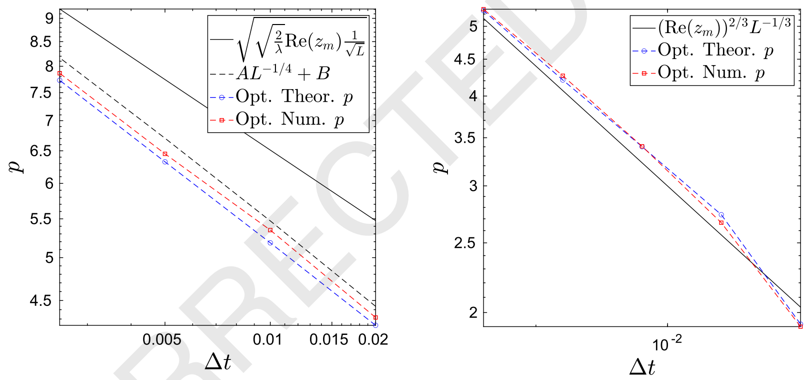


Fig. 6. Left: Asymptotic behavior of the optimal p for overlap $L \approx \Delta t$. ($\tau = 10^{-13}$) **Right:** Asymptotic behavior of the optimal p for overlap $L \approx \Delta t^{\frac{1}{2}}$. ($\tau = 10^{-13}$).

error in terms of Robin interface traces (for both state and adjoint variables). The decay of the error is compared to the theoretical slope of Theorem 7. As expected, the theoretical slope of the decay of the error is asymptotically the same as the numerically measure errors. The theoretical slope is obtained by plotting[2] $3(\max_\kappa \rho_S(\kappa))^{n/2}$. Notice that the non-monotonicity of the numerical errors is due to the complex structure of the spectrum of the iteration matrix of the OSWRM and still consistent with the theoretical bound proved in Theorem 7.

In the last numerical tests, we verify the asymptotic behavior of the optimal parameter p according to the two different choice of the overlap $L \approx \Delta t$ and $L \approx \sqrt{\Delta t}$ and demonstrate the validity of the asymptotic formulas obtained in Theorem 9. In these tests we set $\Delta x = 0.0025$.[3] For this purpose, in Fig. 6 we show the optimal parameter as a function of Δt (black lines) computed using the discrete formulas of Theorem 9. These curves are compared with the value of the optimal p obtained by numerically solving the inf-sup problem (23) (blue line), and the optimal parameter obtained as the one computed by running the OSWRM for different parameters p and finding the one that minimizes the number of iterations needed to make the error smaller than $\tau = 10^{-13}$ (red lines). Notice the great agreement with the three curves for the case $L \approx \Delta t^{\frac{1}{2}}$. However, in the case $L \approx \Delta t$ there is a gap between the asymptotic optimal parameter and the blue and red curves. This behavior is due to the fact that our numerical simulations, even

[2] Note that in Theorem 7 the bound is given for squared norm of the errors at step $2n$, thus the exponent $n/2$ has to be considered in this test. The scaling factor 3 is only for graphical purpose.

[3] We point out that, in this test, the spatial grid is chosen in a way that the discretization error in space is smaller than the one in time.

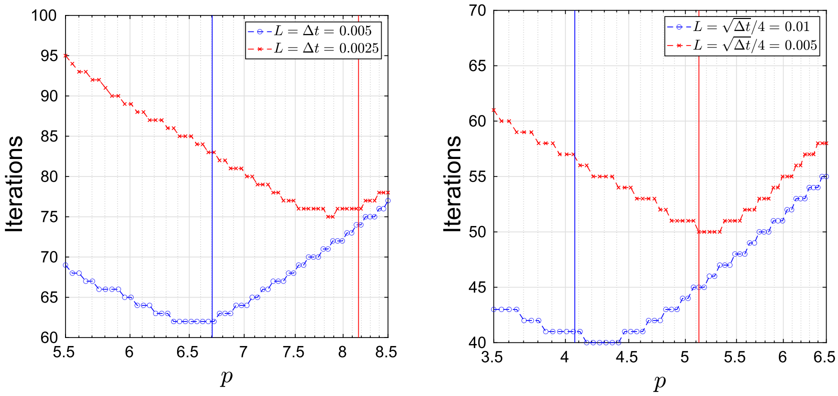


Fig. 7. Number or iterations needed to make the (relative) error smaller than $\tau = 10^{-13}$ (red and blue lines) and values of the theoretical asymptotical optimal parameter (vertical lines). **Left:** Case $L \approx \Delta t$. **Right:** Case $L \approx \Delta t^{\frac{1}{2}}$.

if run for very small Δt, they did not fully reach the asymptotic regime, and a much smaller Δt would be necessary. To rigorously prove this behavior, in Theorem 19 we compute again the optimal parameter p, but this time consider one more term so that p has the form $p = AL^{-\frac{1}{4}} + B$. This is exactly the black-dashed line of Fig. 6 (left), which is now very close to the red and blue lines. This is due to the constant B. In fact, as we are going to show in Theorem 19, the constant A is exactly equal to the one of the optimal p obtained in Theorem 9, while the additional constant B allows us to compensate the gap. Notice that the proof of Theorem 19 is given in the Appendix.

Theorem 19. *In the same settings of Theorem 9, it holds that*

$$p_0^* \sim A\Delta t^{-\frac{1}{4}} + B,$$

where $A = \sqrt{\sqrt{\frac{2}{\lambda}}\mathrm{Re}\left(z_m\right)}$ *and* $B = \frac{A^2(2^{3/2}-4A^2\lambda^{3/2})+8\sqrt{\lambda}\mathrm{Re}(z_m^2)}{8\sqrt{\lambda}\mathrm{Re}(z_m)-\lambda A^2 2^{7/2}}$.

To further demonstrate the validity of our results, we show in Fig. 7 the number of iterations required to reach the relative tolerance $\tau = 10^{-13}$ for varying values of the parameter p and compare these with the theoretical optimal values (vertical lines) obtained by the formulas of Theorem 9. In particular, in the left panel we consider two cases corresponding to $L = \Delta t = 0.005$ (blue curve) and $L = \Delta t = 0.0025$ (red curve). In the right panel, we consider the cases $L = \sqrt{\Delta t}/4 = 0.01$ (blue curve) and $L = \sqrt{\Delta t}/4 = 0.005$ (red curve). In all cases, the theoretical predictions (vertical lines) are very close to the numerical optimum.

6. Conclusion

A convergence analysis for the OSWRM applied to the optimality system of a diffusion-reaction optimization problem with boundary conditions periodic in time was performed. New convergence results were obtained by a semidiscrete Fourier analysis, which allowed the computation of the optimal Robin parameter in both non-overlapping and overlapping cases. Results of numerical experiments confirmed the theoretical findings.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

Gabriele Ciaramella is member of GNCS (Gruppo Nazionale per il Calcolo Scientifico) of INdAM.

Appendix

Proof of Theorem 5. [4] Note that (9c) admits a unique weak solution $\{(Y(U))_s\}_{s=0}^S$ for any given sequence of controls $U \in L_j^2$ for $s = 1, \dots, S$. Let $\mathsf{S}_h : L_1^2 \to H_j^1$ be the linear solution operator associated to (9c), where $H_1^1 := H^1(-\infty, L)$ and $H_2^1 := H^1(0, +\infty)$. We can define the semidiscrete reduced cost functional

$$\widehat{J}_h(U) := \frac{1}{2}\sum_{s=1}^{S} \|(\mathsf{S}_h U)_s - (Y_Q)_s\|_{L_j^2}^2 + \frac{\sigma}{2}\|U_s\|_{L^2}^2 - \lambda(\widetilde{g}_{\mathrm{d},x_j})_s(\mathsf{S}_h U)_s(L).$$

Since the cost function $\widehat{J}_h$ is Fréchet differentiable and strictly convex in U, the first-order necessary and sufficient optimality condition is $(\widehat{J}_h'(\bar{U}))_s = 0$; see, e.g., [35]. Observe that

$$\begin{aligned}&\sum_{s=1}^{S}\langle(\widehat{J}_h'(U))_s, (U^\delta)_s\rangle_{L_j^2} = \sum_{s=1}^{S}\langle(\mathsf{S}_h U)_s - (Y_Q)_s, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2}\\ &+ \sigma\langle U_s, U_s^\delta\rangle_{L_j^2} - \lambda(\widetilde{g}_{\mathrm{d},x_j})_s(\mathsf{S}_h' U^\delta)_s(x_j),\end{aligned}$$

for any sequences U, U^δ with $U_s, U_s^\delta \in L_j^2$. Note that $\mathsf{S}_h' U^\delta$ solves

$$\frac{1}{\Delta t}\langle (Y)_s - (Y)_{s-1}, \varphi\rangle_{L^2} + \lambda\langle (Y_x)_s, \varphi_x\rangle_{L^2} + d\langle (Y)_s, \varphi\rangle_{L^2} + p(Y)_s(L)\varphi(L) = \langle (U^\delta)_s, \varphi\rangle_{L^2} \tag{54}$$

for each $\varphi \in H_j^1$ and $s = 1, \dots, S$ and $(\mathsf{S}_h' U^\delta)_0(x) = (\mathsf{S}_h' U^\delta)_S(x)$. Now, let $Q = \{Q_s\}_{s=0}^S$ with $Q_s \in H_j^1$ the (weak) solution of the equation

$$\begin{aligned}&\frac{1}{\Delta t}\left((Q)_s - (Q)_{s+1}\right) = \lambda(Q_{xx})_s - d(Q)_s + (Y_Q)_s - (Y)_s,\\ &(\partial_{\boldsymbol{n}_j} Q)_s(x_j) + p(Q)_s(x_j) = (\widetilde{g}_{\mathrm{d},x_j})_s,\\ &(Q)_S = (Q)_0.\end{aligned}$$

We have then that

$$\begin{aligned}&\frac{1}{\Delta t}\langle (Q)_s - (Q)_{s+1}, \varphi\rangle_{L_j^2} + \lambda\langle (Q_x)_s, \varphi_x\rangle_{L_j^2} + d\langle (Q)_s, \varphi\rangle_{L_j^2} + p(Q)_s(L)\varphi(L)\\ &= \langle (Y_Q)_s - (\mathsf{S}_h U)_s, \varphi\rangle_{L_j^2} + \lambda(\widetilde{g}_{\mathrm{d},x_j})_s\varphi(x_j)\end{aligned}$$

for each $\varphi \in H_j^1$, $s = 0, \dots, S-1$ and $(Q)_S(x) = (Q)_0(x)$. Now, we can choose $\varphi = (\mathsf{S}_h' U^\delta)_s$ to obtain

$$\begin{aligned}&\frac{1}{\Delta t}\langle (Q)_s - (Q)_{s+1}, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2} + \lambda\langle (Q_x)_s, ((\mathsf{S}_h' U^\delta)_x)_s\rangle_{L_j^2} + d\langle (Q)_s, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2}\\ &+ p(Q)_s(x_j)(\mathsf{S}_h' U^\delta)_s(x_j) = \langle (Y_Q)_s - (\mathsf{S}_h U)_s, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2} + \lambda(\widetilde{g}_{\mathrm{d},x_j})_s(\mathsf{S}_h' U^\delta)_s(x_j)\end{aligned}$$

which leads to

$$\begin{aligned}\langle \widehat{J}'(U), U^\delta\rangle_{X_1} &= \sum_{s=0}^{S-1}\frac{1}{\Delta t}\langle (Q)_{s+1} - (Q)_s, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2} - \lambda\langle (Q_x)_s, ((\mathsf{S}_h' U^\delta)_x)_s\rangle_{L_j^2}\\ &\quad - d\langle (Q)_s, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2} - p(Q)_s(x_j)(\mathsf{S}_h' U^\delta)_s(x_j) + \sigma\sum_{s=1}^{S}\langle (U)_s, (U^\delta)_s\rangle_{L_j^2}\\ &= \sum_{s=0}^{S-1}\frac{1}{\Delta t}\langle (\mathsf{S}_h' U^\delta)_s, (Q)_{s+1}\rangle_{L_j^2} + \sum_{s=1}^{S} -\frac{1}{\Delta t}\langle (\mathsf{S}_h' U^\delta)_s, (Q)_s\rangle_{L_j^2} - d\langle (\mathsf{S}_h' U^\delta)_s, (Q)_s\rangle_{L_j^2}\\ &\quad - \lambda\langle ((\mathsf{S}_h' U^\delta)_x)_s, (Q_x)_s\rangle_{L_j^2} - p(\mathsf{S}_h' U^\delta)_s(x_j)(Q)_s(x_j) + \sigma\langle (U)_s, (U^\delta)_s\rangle_{L_j^2}\\ &= \cancel{\sum_{s=0}^{S-1}\frac{1}{\Delta t}\langle\langle (Q)_{s+1}, (\mathsf{S}_h' U^\delta)_s\rangle_{L_j^2}} - \cancel{\sum_{s=1}^{S}\frac{1}{\Delta t}\langle (Q)_s, (\mathsf{S}_h' U^\delta)_{s-1}\rangle_{L_j^2}})\\ &\quad + \sum_{s=1}^{S}\langle -(Q)_s + \sigma(U)_s, (U^\delta)_s\rangle_{L_j^2},\end{aligned}$$

where we used the periodic conditions and (54) tested for $\varphi = (Q)_s \in H_j^1$ for $s = 1, \dots, S$. This shows that $(\nabla J(U))_s = \langle -(Q)_s + \sigma(U)_s, (U^\delta)_s\rangle_{L_j^2}$ for each $\{(U)\}_{s=1}^S$ and thus $0 = (\nabla J(\bar{U}))_s = -(Q(\bar{U}))_s + \sigma(\bar{U})_s$. This means that $(\bar{U})_s = \sigma^{-1}(Q(\bar{U}))_s$ and thus the first-order necessary and sufficient optimality system of the problem of minimizing (10) subject to (9c) can be expressed as (9). □

[4] One can equivalently prove it also by using the Lagrange formalism and imposing the periodic condition $Q_j(0,\cdot) = Q_j(S,\cdot)$ for the Lagrange multiplier Q_j.

Proof of Theorem 19. The proof is exactly the one of Section 4.3.1. Only Step 3 needs to be recomputed. To obtain the coefficients A and B, we use the ansatz $p = AL^{-\frac{1}{4}} + B$, notice that

$$\frac{p}{z_M} = \sqrt{\frac{\lambda}{2}}AL^{1/4} + \sqrt{\frac{\lambda}{2}}AL^{1/2} + O(L^{5/4}),$$

and consider more terms in the expansion of $\frac{|z-1|^2}{|z+1|^2}$:

$$\left|\frac{1-z}{1+z}\right|^2 = 1 - 4\mathrm{Re}\,(z) + 8\mathrm{Re}\left(z^2\right) - 12\mathrm{Re}\left(z^3\right) + O(z^4). \tag{55}$$

Since $\frac{p}{z_M} \ll 1$ for $L \ll 1$, we can use (55) for $z = \frac{p}{z_M}$ and recall that $e^{-2L\mathrm{Re}(z_M)} \sim 1 - 2L\sqrt{2a} = 1 - 2\sqrt{\frac{2}{\lambda}}L^{1/2} + O(L)$, to obtain

$$R(z_M, p, L) = 1 - 4\sqrt{\frac{2}{\lambda}}L^{1/4} - 2\Big(2\sqrt{\frac{\lambda}{2}}B - 2\lambda A^2 + \sqrt{\frac{2}{\lambda}}\Big)L^{1/2} + 8A\Big(\lambda B + 1\Big)L^{3/4} + O(L). \tag{56}$$

Proceeding in a similar way, we obtain

$$R(z_m, p, L) = 1 - 4\frac{\mathrm{Re}\left(z_m\right)}{p} + 8\frac{\mathrm{Re}\left(z_m^2\right)}{p^2} - 12\frac{\mathrm{Re}\left(z_m^3\right)}{p^3} + O\Big(\frac{z_m^4}{p^4}\Big). \tag{57}$$

Using (56) and (57), we can compute the expansion

$$R(z_m, p, L) - R(z_M, p, L) = -\frac{1}{BL^{1/4} + A}[C_1 L^{1/4} + C_2 L^{2/4} + O(L^{3/4})], \tag{58}$$

where

$$C_1 = A^2(4\sqrt{\lambda}\mathrm{Re}\left(z_m\right) - 2^{3/2}A^2\lambda),$$
$$C_2 = A(-8\sqrt{\lambda}\mathrm{Re}\left(z_m^2\right) + 8B\sqrt{\lambda}\mathrm{Re}\left(z_m\right) + 4A^4\lambda^{3/2} - 2^{7/2}A^2B\lambda - 2^{3/2}A^2).$$

Now, the result follows by setting to zero the two higher order terms of (58), that is setting $C_1 = 0$ and $C_2 = 0$. □

References

[1] O. Axelsson, Z.-Z. Liang, A note on preconditioning methods for time-periodic eddy current optimal control problems, J. Comput. Appl. Math. 352 (2019) 262–277.
[2] O. Axelsson, D. Lukáš, Preconditioning methods for eddy-current optimally controlled time-harmonic electromagnetic problems, J. Numer. Math. 27 (1) (2019) 1–21.
[3] M. Kolmbauer, The multiharmonic finite element and boundary element method for simulation and control of eddy current problems, PhD Thesis.
[4] M. Kolmbauer, U. Langer, A robust preconditioned MinRes solver for distributed time-periodic eddy current optimal control problems, SIAM J. Sci. Comput. 34 (6) (2012) 785–809.
[5] O. Axelsson, S. Farouq, M. Neytcheva, A preconditioner for optimal control problems, constrained by Stokes equation with a time-harmonic control, J. Comput. Appl. Math. 310 (2017) 5–18.
[6] B. Houska, M. Diehl, Robustness and stability optimization of power generating kite systems in a periodic pumping mode, in: 2010 IEEE International Conference on Control Applications, 2010, pp. 2172–2177.
[7] B. Houska, F. Logist, J. Van Impe, M. Diehl, Approximate robust optimization of time-periodic stationary states with application to biochemical processes, in: Proceedings of the 48h IEEE Conference on Decision and Control, 2009, pp. 6280–6285.
[8] F. Logist, A.V. Wouwer, I. Smets, J. Van Impe, Optimal temperature profiles for tubular reactors implemented through a flow reversal strategy, Chem. Eng. Sci. 62 (17) (2007) 4675–4688.
[9] M. Gunzburger, C. Trenchea, Optimal control of the time-periodic MHD equations, Nonlinear Anal. 63 (5–7) (2005) 1687–1699.
[10] M. Gunzburger, C. Trenchea, Analysis and discretization of an optimal control problem for the time-periodic MHD equations, J. Math. Anal. Appl. 308 (2) (2005) 440–466.
[11] M. Kollmann, M. Kolmbauer, U. Langer, M. Wolfmayr, W. Zulehner, A robust finite element solver for a multiharmonic parabolic optimal control problem 65 (3) (2013).
[12] D. Abbeloos, M. Diehl, M. Hinze, S. Vandewalle, Nested multigrid methods for time-periodic, parabolic optimal control problems, Comput. Vis. Sci. 14 (1) (2011) 27–38.
[13] C.V. Pao, Numerical methods for time-periodic solutions of nonlinear parabolic boundary value problems, SIAM J. Numer. Anal. 39 (2) (2002) 647–667.
[14] S. Vandewalle, R. Piessens, On dynamic iteration methods for solving time-periodic differential equations, SIAM J. Numer. Anal. 30 (1) (1993) 286–303.
[15] W. Hackbusch, Fast numerical solution of time-periodic parabolic problems by a multigrid method, SIAM J. Sci. Comput. 2 (2) (1981) 198–206.
[16] A. Bensoussan, R. Glowinski, J.-L. Lions, Méthode de décomposition appliquée au contrôle optimal de systemes distribués, in: IFIP Technical Conference on Optimization Techniques, Springer, 1973, pp. 141–151.
[17] J.E. Lagnese, G. Leugering, Domain in Decomposition Methods in Optimal Control of Partial Differential Equations, Springer, New York, 2004.
[18] M. Heinkenschloss, H. Nguyen, Domain decomposition preconditioners for linear-quadratic elliptic optimal control problems, CAAM Technical Reports, https://hdl.handle.net/1911/102032, 2004.
[19] B. Delourme, L. Halpern, A complex homographic best approximation problem. Application to optimized Robin–Schwarz algorithms, and optimal control problems, SIAM J. Numer. Anal. 59 (3) (2021) 1769–1810.
[20] J.-D. Benamou, A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations, SIAM J. Numer. Anal. 33 (6) (1996) 2401–2416.
[21] J.-D. Benamou, B. Desprès, A domain decomposition method for the Helmholtz equation and related optimal control problems, Research Report RR-2791, INRIA, 1996.

[22] R. Bartlett, M. Heinkenschloss, D. Ridzal, B. van Bloemen Waanders, Domain decomposition methods for advection dominated linear quadratic elliptic optimal control problem, Comput. Methods Appl. Mech. Eng. 195 (2006).
[23] M. Heinkenschloss, H. Nguyen, Balancing Neumann-Neumann methods for elliptic optimal control problems, in: Domain Decomposition Methods in Science and Engineering, Springer, 2005, pp. 589–596.
[24] M.J. Gander, F. Kwok, B.C. Mandal, Convergence of substructuring methods for elliptic optimal control problems, in: Domain Decomposition Methods in Science and Engineering XXIV, Springer, 2018, pp. 291–300.
[25] J.-D. Benamou, A domain decomposition method for control problems, in: Proceedings of the 9th International Conference on Domain Decomposition Methods, 1998, pp. 266–273.
[26] M.J. Gander, A.M. Stuart, Space-time continuous analysis of waveform relaxation for the heat equation, SIAM J. Sci. Comput. 19 (6) (1998) 2014–2031.
[27] M.J. Gander, L. Halpern, Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems, SIAM J. Numer. Anal. 45 (2) (2007) 666–697.
[28] D. Bennequin, M.J. Gander, L. Gouarin, L. Halpern, Optimized Schwarz waveform relaxation for advection reaction diffusion equations in two dimensions, Numer. Math. 134 (3) (2016) 513–567.
[29] M.J. Gander, F. Kwok, B.C. Mandal, Dirichlet-Neumann waveform relaxation methods for parabolic and hyperbolic problems in multiple subdomains, BIT Numer. Math. 1 (2021) 1–35.
[30] M.J. Gander, L. Halpern, F. Hubert, S. Krell, Discrete optimization of Robin transmission conditions for anisotropic diffusion with discrete duality finite volume methods, Vietnam J. Math. 49 (4) (2021) 1349–1378.
[31] D. Bennequin, M.J. Gander, L. Halpern, A homographic best approximation problem with application to optimized Schwarz waveform relaxation, Math. Comput. 78 (2009) 185–223.
[32] J.L. Lions, E. Magenes, Non-homogeneous Boundary Value Problems and Applications (vol. I), Die Grundlehren der Mathematischen Wissenschaften, Springer, Berlin Heidelberg, 1972.
[33] P.G. Ciarlet, Introduction to Numerical Linear Algebra and Optimisation, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 1989.
[34] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen, Springer, Berlin Heidelberg, 1971.
[35] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods, and Applications, Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, 2010.
[36] M.J. Gander, Optimized Schwarz methods, SIAM J. Numer. Anal. 44 (2) (2006) 699–731.
[37] F. Chaouqui, G. Ciaramella, M.J. Gander, T. Vanzan, On the scalability of classical one-level domain-decomposition methods, Vietnam J. Math. 46 (4) (2018) 1053–1088.
[38] J.L. Lions, E. Magenes, Non-homogeneous Boundary Value Problems and Applications (vol. II), Die Grundlehren der Mathematischen Wissenschaften, Springer, Berlin Heidelberg, 1972.
[39] J.E. Lagnese, G. Leugering, Time-domain decomposition of optimal control problems for the wave equation, Syst. Control Lett. 48 (3) (2003) 229–242.
[40] M. Heinkenschloss, H. Nguyen, Neumann-Neumann domain decomposition preconditioners for linear-quadratic elliptic optimal control problems, SIAM J. Sci. Comput. 28 (3) (2006) 1001–1028.
[41] G. Ciaramella, M.J. Gander, Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: part I, SIAM J. Numer. Anal. 55 (3) (2017) 1330–1356.
[42] G. Ciaramella, M.J. Gander, P. Mamooler, How to best choose the outer coarse mesh in the domain decomposition method of Bank and Jimack, Vietnam J. Math. (2022) 1–33.
[43] S. Loisel, J. Côté, M.J. Gander, L. Laayouni, A. Qaddouri, Optimized domain decomposition methods for the spherical Laplacian, SIAM J. Numer. Anal. 48 (2) (2010) 524–551.
[44] A. Arnoult, C. Japhet, P. Omnes, Discrete-time analysis of Schwarz waveform relaxation convergence, working paper or preprint, 2022, https://sorbonne-paris-nord.hal.science/hal-03746438.
[45] M.J. Gander, L. Halpern, F. Hubert, S. Krell, Optimized Robin transmission conditions for anisotropic diffusion on arbitrary meshes, in: XXVII International Symposium on Domain Decomposition Methods for Partial Differential Equations, 2023, in press.
[46] G. Cardano, T.R. Witmer, O. Ore, The Rules of Algebra: Ars Magna. Alianza Universidad, Dover Publications, Dover, 2007.

Highlights

- Treatment of Robin-type domain decomposition methods for parabolic/periodic optimal control problems.
- Semidiscrete convergence analysis.
- Optimization of the Robin parameter in asymptotic regime.