Optimized Schwarz Method for Coupled Direct-Adjoint Problems Applied to Parameter Identification in Advection-Diffusion Equation

Alexandre Vieira and Pierre-Henri Cocquet

1 Introduction

Let Ω be a bounded open set with Lipschitz boundary. We want to solve the problem

$$\begin{aligned} &\min_k \frac{1}{2}\|T - T_{\text{target}}\|^2_{L^2(\Omega)}, \\ &\text{s.t. } \operatorname{div}(\mathbf{u}T) - \operatorname{div}(k\nabla T) = f \text{ in } \Omega,\ T|_{\partial\Omega} = T_0 \text{ and } k \in U_{ad}, \end{aligned} \tag{1}$$

with $f \in L^2(\Omega), \mathbf{u} \in H^1(\Omega)$ given s.t. $\operatorname{div}\mathbf{u} = 0$. The set of admissible control U_{ad} contains all $k(x) \in [a, b]$ for a.e. x and $a > 0$ and is chosen such that any sequences $(k_n)_n \subset U_{ad}$ have a subsequence converging a.e. in Ω. Such $k \in U_{ad}$ must be more regular (e.g with bounded variation) and we refer to [9, p. 9, Assumption 1] for an example of such U_{ad}. With all these assumptions, (1) has at least one optimal solution (see e.g. [9, Theorem 3]).

We will focus on finding ways to compute a solution to (1) on several subdomains. In recent years, a lot of papers started to look at ways to decompose the resolution of optimal control problems. In [5], the authors split the optimization problem as two independent optimization problems, splitted by subdomains, with an augmented cost. The necessary (and sufficient) conditions of optimality let us see that it actually reduces to a classical Schwarz method applied to the direct and adjoint systems, where the control could be eliminated (see also [1, 2, 7]).

These papers rely on the huge literature analyzing the different flavors of the Schwarz iterative method: we only refer to [4] for an introduction and to [8] for a more in depth presentation of these methods (and other decomposition methods).

A. Vieira
Université de la Réunion, PIMENT, Sainte-Clotilde, France,
e-mail: alexandre.vieira@univ-reunion.fr

P.-H. Cocquet
Université de Pau et des Pays de l'Adour, SIAME, rue de l'Université, Pau, France,
e-mail: pierre-henri.cocquet@univ-pau.fr

In order to decompose the resolution of (1) across several subdomains, we will also adopt an *indirect approach*: we will decompose the computations of the gradient of the cost, which will be used afterward in a descent method. Since the control needs to be defined on the whole domain Ω, it seems hard to define a decomposition of (1) using an overlap. Therefore, we will focus on finding an optimized Schwarz iteration, without overlap, to compute the gradient of the cost.

2 Direct and adjoint equations

First of all, we express the gradient of the cost which can be given thanks to an adjoint equation. The next result can be proved using [6, Corollary 1.3].

Theorem 1 *Let* $J(k) = \frac{1}{2}\|T(k) - T_{target}\|^2_{L^2(\Omega)}$ *with* $T(k) \in H^1(\Omega)$ *be the solution to* $\operatorname{div}(\mathbf{u}T) - \operatorname{div}(k\nabla T) = f$ *in* Ω *and* $T|_{\partial\Omega} = T_0$*. Then:* $\partial_k J(k) = \nabla T \cdot \nabla\lambda$*, where* λ *solves*

$$\operatorname{div}(k\nabla\lambda - \mathbf{u}\lambda) = T - T_{target},\ \lambda|_{\partial\Omega} = 0.$$

Therefore, the gradient of J can be computed by solving for fixed k:

$$\begin{cases} -\operatorname{div}(k\nabla T - \mathbf{u}T) = f, & T|_{\partial\Omega} = T_0, \\ \ \ \operatorname{div}(k\nabla\lambda + \mathbf{u}\lambda) = T - T_{\text{target}}, & \lambda|_{\partial\Omega} = 0. \end{cases} \tag{2}$$

We now expose our strategy in order to accelerate the resolution of (1): we would like to decompose the resolution of (2) across several subdomains in order to accelerate the computation of the gradient. However, since our optimization parameter k is defined on Ω, it seems hard to imagine a decomposition method using an overlap. If we were computing a solution of (2) on two subdomains Ω_1 and Ω_2 with an overlap, then we would end with two different gradients $\partial_k J(k)$ on $\Omega_1 \cap \Omega_2$, depending on which side the gradient is computed. Therefore, using a descent technique would produce two different controls k_1 on Ω_1 and k_2 on Ω_2, with possibly different values on $\Omega_1 \cap \Omega_2$. This could prevent the convergence to an optimal solution of (1).

To summarize, we are interested in non-overlapping Schwarz techniques. It should be noted that optimized Schwarz method for an advection-diffusion equation has been done in [3] but, to the best of our knowledge, never for (2).

3 Optimized Schwarz method for coupled direct-adjoint system

We assume there is open sets Ω_i such that $\overline{\Omega} = \overline{\Omega_1 \cup \Omega_2}$ with interface $\Gamma_\cap = \overline{\Omega_1} \cap \overline{\Omega_2}$. A non-overlapping Schwarz method for (2) can then be roughly defined as

1. Take an initial guess (T_i^0, λ_i^0) defined on Ω_i,
2. Until some stopping criteria are met: Compute $(T_i^{n+1}, \lambda_i^{n+1})$ satisfying

$$\begin{cases} -\mathrm{div}(k\nabla T_i^{n+1}) + \mathrm{div}(uT_i^{n+1}) = f, & \text{on } \Omega_i,\ T_i^{n+1}|_{\partial\Omega_i\setminus\Gamma_\cap} = T_0, \\ \mathrm{div}(k\nabla\lambda_i^{n+1}) + \mathrm{div}(u\lambda_i^{n+1}) = T^{n+1} - T_{\text{target}} & \text{on } \Omega_i,\ \lambda_i^{n+1}|_{\partial\Omega_i\setminus\Gamma_\cap} = 0, \end{cases} \tag{3}$$

and the following transmission conditions on the interface

$$\begin{aligned} k\partial_{\mathbf{n}}\begin{pmatrix} T_i^{n+1} \\ \lambda_i^{n+1} \end{pmatrix} - \frac{\mathbf{u}\cdot\mathbf{n}}{2}\begin{pmatrix} T_i^{n+1} \\ -\lambda_i^{n+1} \end{pmatrix} + (-1)^{i+1}\mathcal{S}_i\begin{pmatrix} T_i^{n+1}|_{\Gamma_\cap} \\ \lambda_i^{n+1}|_{\Gamma_\cap} \end{pmatrix} \\ = k\partial_{\mathbf{n}}\begin{pmatrix} T_{3-i}^{n} \\ \lambda_{3-i}^{n} \end{pmatrix} - \frac{\mathbf{u}\cdot\mathbf{n}}{2}\begin{pmatrix} T_{3-i}^{n} \\ -\lambda_{3-i}^{n} \end{pmatrix} + (-1)^{i+1}\mathcal{S}_i\begin{pmatrix} T_{3-i}^{n}|_{\Gamma_\cap} \\ \lambda_{3-i}^{n}|_{\Gamma_\cap} \end{pmatrix}, \end{aligned} \tag{4}$$

where $\mathbf{n}$ is the outer normal to $\partial\Omega_1$. In (3)–(4), $\mathcal{S}_i$ are linear operators acting on traces of (T_i, λ_i) (e.g. $\mathcal{S}_i = p_i$ id where p_i are some constants and id is the identity operator, or some linear differential operator involving tangential derivatives).

To study the convergence of the non-overlapping Schwarz method as well as its convergence properties, we are going to restrict ourselves to the case k = constant, $\Omega = \mathbb{R}^2$, $\Omega_1 = (-\infty, 0) \times \mathbb{R}$ and $\Omega_2 = (0, +\infty) \times \mathbb{R}$. In such setting, we can rely on Fourier analysis [4] to study the convergence and also to design optimized transmission operators $\mathcal{S}_i$ that accelerate the convergence. Without loss of generality, we also suppose that $f = T_{\text{target}} = 0$ since we are interested in the error.

3.1 Computation of the optimal transmission operator

We start by applying Fourier transform to (2) along the y axis:

$$-k\partial_{xx}\begin{pmatrix} \hat{T}_i^n \\ \hat{\lambda}_i^n \end{pmatrix} + u_1\partial_x\begin{pmatrix} \hat{T}_i^n \\ -\hat{\lambda}_i^n \end{pmatrix} + \begin{pmatrix} k\omega^2 - iu_2\omega & 0 \\ 1 & k\omega^2 + iu_2\omega \end{pmatrix}\begin{pmatrix} \hat{T}_i^n \\ \hat{\lambda}_i^n \end{pmatrix} = 0, i \in \{1, 2\}.$$

Along x, this is a second order ordinary differential equation which can be solved explicitly. Define :

$$r_\pm^T(\omega) = \frac{u_1 \pm \sqrt{u_1^2 + 4k^2\omega^2 - 4iku_2\omega}}{2k},\ r_\pm^\lambda(\omega) = \frac{-u_1 \pm \sqrt{u_1^2 + 4k^2\omega^2 + 4iku_2\omega}}{2k}.$$

Concerning $\hat{T}$, using the Dirichlet condition at infinity, there exist functions $A_T^n(\omega)$ and $B_T^n(\omega)$ such that:

$$\hat{T}_1^n(x,\omega) = A_T^n(\omega)e^{r_+^T(\omega)x},\ \hat{T}_2^n(x,\omega) = B_T^n(\omega)e^{r_-^T(\omega)x}.$$

These solutions are reintroduced into the equation in order to solve it for $\hat{\lambda}$. There, the equation is non-homogeneous, but the right hand-side is of the form $C(\omega)e^{D(\omega)x}$, for some functions C and D independent of x. An arbitrary solution is therefore easily found, and one proves that they take the form:

$$\hat{\lambda}_1^n(x,\omega) = A_\lambda^n(\omega)e^{r_+^\lambda(\omega)x} - A_{\lambda T}(\omega)\hat{T}_1^n(x,\omega),$$
$$\hat{\lambda}_2^n(x,\omega) = B_\lambda^n(\omega)e^{r_-^\lambda(\omega)x} - B_{\lambda T}(\omega)\hat{T}_2^n(x,\omega),$$

where $A_{\lambda T}(\omega) = (-kr_+^T(\omega)^2 - u_1 r_+^T(\omega) + k\omega^2 + iu_2\omega)^{-1}$ and $B_{\lambda T}(\omega) = (-kr_-^T(\omega)^2 - u_1 r_-^T(\omega) + k\omega^2 + iu_2\omega)^{-1}$.

We may now derive each solution with x:

$$\partial_x \hat{T}_1^n(x,\omega) = r_+^T(\omega)\hat{T}_1^n(x,\omega),\ \partial_x \hat{T}_2^n(x,\omega) = r_-^T(\omega)\hat{T}_2^n(x,\omega),$$
$$\partial_x \hat{\lambda}_1^n(x,\omega) = r_+^\lambda(\omega)A_\lambda^n(\omega)e^{r_+^\lambda(\omega)x} - A_{\lambda T}(\omega)r_+^T(\omega)\hat{T}_1^n(x,\omega),$$
$$\partial_x \hat{\lambda}_2^n(x,\omega) = r_-^\lambda(\omega)B_\lambda^n(\omega)e^{r_-^\lambda(\omega)x} - B_{\lambda T}(\omega)r_-^T(\omega)\hat{T}_2^n(x,\omega).$$

We now assume that $\mathcal{F}_y(\mathcal{S}_i(T,\lambda))(x,\omega) = \sigma_i(\omega)(\hat{T},\hat{\lambda})$, where σ_i is a 2×2 complex matrix. The transmission conditions then read:

$$\begin{aligned}
k\partial_x \begin{pmatrix}\hat{T}_1^n\\ \hat{\lambda}_1^n\end{pmatrix}(x,\omega) - \frac{u_1}{2}\begin{pmatrix}\hat{T}_1^n\\ -\hat{\lambda}_1^n\end{pmatrix}(x,\omega) + \sigma_1(\omega)\begin{pmatrix}\hat{T}_1^n\\ \hat{\lambda}_1^n\end{pmatrix}(x,\omega)&\\
= \left(M_r^+(x,\omega) + \sigma_1(\omega)M_0^+(x,\omega)\right)\begin{pmatrix}A_T^n(\omega)\\ A_\lambda^n(\omega)\end{pmatrix}&,\\
k\partial_x \begin{pmatrix}\hat{T}_2^n\\ \hat{\lambda}_2^n\end{pmatrix}(x,\omega) - \frac{u_1}{2}\begin{pmatrix}\hat{T}_2^n\\ -\hat{\lambda}_2^n\end{pmatrix}(x,\omega) + \sigma_2(\omega)\begin{pmatrix}\hat{T}_2^n\\ \hat{\lambda}_2^n\end{pmatrix}(x,\omega)&\\
= \left(M_r^-(x,\omega) + \sigma_2(\omega)M_0^-(x,\omega)\right)\begin{pmatrix}B_T^n(\omega)\\ B_\lambda^n(\omega)\end{pmatrix}&,
\end{aligned}$$

$$M_r^+(x,\omega) = \begin{pmatrix}(kr_+^T(\omega) - \frac{u_1}{2})e^{r_+^T(\omega)x} & 0\\ -(kr_+^T(\omega) - \frac{u_1}{2})A_{\lambda T}(\omega)e^{r_+^T(\omega)x} & (kr_+^\lambda(\omega) - \frac{u_1}{2})e^{r_+^\lambda(\omega)x}\end{pmatrix}$$

$$M_0^+(x,\omega) = \begin{pmatrix}e^{r_+^T(\omega)x} & 0\\ A_{\lambda T}(\omega)e^{r_+^T(\omega)x} & e^{r_+^\lambda(\omega)x}\end{pmatrix},\ M_0^-(x,\omega) = \begin{pmatrix}e^{r_-^T(\omega)x} & 0\\ -B_{\lambda T}(\omega)e^{r_-^T(\omega)x} & e^{r_-^\lambda(\omega)x}\end{pmatrix}$$

$$M_r^-(x,\omega) = \begin{pmatrix}(kr_-^T(\omega) - \frac{u_1}{2})e^{r_-^T(\omega)x} & 0\\ -(kr_-^T(\omega) - \frac{u_1}{2})B_{\lambda T}(\omega)e^{r_-^T(\omega)x} & (kr_-^\lambda(\omega) - \frac{u_1}{2})e^{r_-^\lambda(\omega)x}\end{pmatrix}$$

Using the conditions at $x = 0$, we get the following recurrence:

$$\begin{aligned}
\begin{pmatrix}A_T^n(\omega)\\ A_\lambda^n(\omega)\end{pmatrix} = &\underbrace{\left[M_r^+(0,\omega) + \sigma_1(\omega)M_0^+(0,\omega)\right]^{-1}\left[M_r^-(0,\omega) + \sigma_1(\omega)M_0^-(0,\omega)\right]}_{M_1(\omega)}\\
&\underbrace{\left[M_r^-(0,\omega) - \sigma_2(\omega)M_0^-(0,\omega)\right]^{-1}\left[M_r^+(0,\omega) - \sigma_2(\omega)M_0^+(0,\omega)\right]}_{M_2(\omega)}\\
&\times\begin{pmatrix}A_T^{n-2}(\omega)\\ A_\lambda^{n-2}(\omega)\end{pmatrix}
\end{aligned}$$

$$\begin{pmatrix} B_T^n(\omega) \\ B_\lambda^n(\omega) \end{pmatrix} = \underbrace{\left[M_r^-(0,\omega) - \sigma_2(\omega) M_0^-(0,\omega)\right]^{-1} \left[M_r^+(0,\omega) - \sigma_2(\omega) M_0^+(0,\omega)\right]}_{M_2(\omega)}$$
$$\underbrace{\left[M_r^+(0,\omega) + \sigma_1(\omega) M_0^+(0,\omega)\right]^{-1} \left[M_r^-(0,\omega) + \sigma_1(\omega) M_0^-(0,\omega)\right]}_{M_1(\omega)}$$
$$\times \begin{pmatrix} B_T^{n-2}(\omega) \\ B_\lambda^{n-2}(\omega) \end{pmatrix}$$

Therefore, the optimal choice of σ_i cancels $M_1(\omega)M_2(\omega)$ and $M_2(\omega)M_1(\omega)$; this reads:

$$\begin{aligned}
\sigma_1^{\mathrm{opt}}(\omega) &= -M_r^-(0,\omega)\left(M_0^-(0,\omega)\right)^{-1} \\
&= \begin{pmatrix} -kr_-^T(\omega) + \frac{u_1}{2} & 0 \\ -kB_{\lambda T}(\omega)[r_-^\lambda(\omega) - r_-^T(\omega)] & -kr_-^\lambda(\omega) - \frac{u_1}{2} \end{pmatrix}, \\
\sigma_2^{\mathrm{opt}}(\omega) &= M_r^+(0,\omega)\left(M_0^+(0,\omega)\right)^{-1} \\
&= \begin{pmatrix} kr_+^T(\omega) - \frac{u_1}{2} & 0 \\ -kA_{\lambda T}(\omega)[r_+^T(\omega) - r_+^\lambda(\omega)] & kr_+^\lambda(\omega) + \frac{u_1}{2} \end{pmatrix}.
\end{aligned}$$

However, as it is usual concerning the optimal Schwarz operator, an inverse Fourier transform proves that $\mathcal{S}_i^{opt}$, the inverse Fourier transform of σ_i^{opt}, is a non-local operator [4]. This property can be difficult to handle in a numerical method. This is why we will restrict the set of admissible transmission operator $\mathcal{S}_i$ to local constant operators.

3.2 Computation of optimized transmission operator

Instead of using the optimal (non-local) operator, we will search for an optimal lower-triangular matrix P_i, which we will suppose to be constant in ω. All the calculations above can be done similarly with this new assumption, and we may write similarly the new matrices $M_1(\omega)$ and $M_2(\omega)$. Suppose $\sigma_1 = \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{13} & \sigma_{14} \end{pmatrix}$ and $\sigma_2 = \begin{pmatrix} \sigma_{21} & 0 \\ \sigma_{23} & \sigma_{24} \end{pmatrix}$. Then $M_l(\omega)M_m(\omega) = \begin{pmatrix} M_1^1(\omega) & 0 \\ M_3^{lm}(\omega) & M_4^1(\omega) \end{pmatrix}$, where

$$M_1^1(\omega) = \frac{(2kr_-^T(\omega) + 2\sigma_{11} - u_1)(-2kr_+^T(\omega) + 2\sigma_{21} + u_1)}{(2kr_+^T(\omega) + 2\sigma_{11} - u_1)(-2kr_-^T(\omega) + 2\sigma_{21} + u_1)},$$

$$M_4^1(\omega) = \frac{(2kr_-^\lambda(\omega) + 2\sigma_{14} + u_1)(2kr_+^\lambda(\omega) - 2\sigma_{24} + u_1)}{(2kr_+^\lambda(\omega) + 2\sigma_{14} + u_1)(2kr_-^\lambda(\omega) - 2\sigma_{24} + u_1)},$$

and $M_3^{lm}(\omega)$ for $l, m = 1, 2$ can be computed as above but are not given since their expressions are not needed in the subsequent analysis.

A way that seems natural is to optimize the transmission conditions consists in solving the min-max problem: $\min_{\sigma_{11},\sigma_{13},\sigma_{14}} \max_{\omega\in[\omega_1,\omega_2]} \|M_1(\omega)M_2(\omega)\|$ for some matrix norm $\|\cdot\|$ and some constants $\omega_1 < \omega_2$. Solving this min-max problem can be tricky: the result may depend on the chosen norm, and the complicated expression of the components of M_1M_2 makes the whole analysis inextricable. Furthermore, it is not entirely clear how one could use either the product M_1M_2 or M_2M_1 in this min max problem. This could change the nature of the result.

However, we remark that the spectral radius appears to be useful in this case, since $\rho(M_1(\omega)M_2(\omega)) = \rho(M_2(\omega)M_1(\omega))$ only depend on σ_{11}, σ_{21}, σ_{14} and σ_{24}. Furthermore, optimizing the spectral radius of the matrices may be done in two independent optimization problems:

$$\min_{\sigma_{11},\sigma_{21}} \max_{\omega\in[\omega_1,\omega_2]} \left| \frac{\left(-\sqrt{u_1^2+4k^2\omega^2-4iku_2\omega}+2\sigma_{11}\right)\left(-\sqrt{u_1^2+4k^2\omega^2-4iku_2\omega}+2\sigma_{21}\right)}{\left(\sqrt{u_1^2+4k^2\omega^2-4iku_2\omega}+2\sigma_{11}\right)\left(\sqrt{u_1^2+4k^2\omega^2-4iku_2\omega}+2\sigma_{21}\right)} \right|,$$

$$\min_{\sigma_{14},\sigma_{24}} \max_{\omega\in[\omega_1,\omega_2]} \left| \frac{\left(-\sqrt{u_1^2+4k^2\omega^2+4iku_2\omega}+2\sigma_{14}\right)\left(-\sqrt{u_1^2+4k^2\omega^2+4iku_2\omega}+2\sigma_{24}\right)}{\left(\sqrt{u_1^2+4k^2\omega^2+4iku_2\omega}+2\sigma_{14}\right)\left(\sqrt{u_1^2+4k^2\omega^2+4iku_2\omega}+2\sigma_{24}\right)} \right|.$$

This kind of min max problem can be solved. Suppose $\omega_1 = 0$, $\omega_2 = \pi/h$ and applying results from [3, p. 35, Eq. (2.11)] prove, assuming h is small enough, that the solution in this case is:

$$\sigma_{11} = \sigma_{14} = \left(\frac{k\pi|u_1|^3}{2h}\right)^{\frac{1}{4}}, \ \sigma_{21} = \sigma_{24} = \left(\frac{2^5k^3\pi^3|u_1|}{h^3}\right)^{\frac{1}{4}}. \tag{5}$$

Other similar results can be found in [3]. However, this approach of optimizing the spectral radius let the parameters σ_{13} and σ_{23} free, and $M_3^{12}(\omega)$ and $M_3^{21}(\omega)$ both depend on σ_{13} and σ_{23}. We show below how these parameters can be chosen.

3.3 Random tests for the last parameters

In order to see the influence of the parameters σ_{i3}, we ran a batch of numerical tests with random values for σ_{i3}. We then solve (2) on $\Omega = (-1,1)\times(0,1)$ with $k = 1$, $f = T_{\text{target}} = 0$, $\mathbf{u} = (-2,0)$, $T_0|_{x=-1} = 0$, $T_0|_{x=1} = 2$ and $T_0|_{y\in\{0,1\}} = 1$. We first solve the equation on Ω, and compare it with the solution of (3)–(4), where $\Omega_1 = (-1,0)\times(0,1)$, $\Omega_2 = (0,1)\times(0,1)$, $\Gamma_\cap = \{0\}\times[0,1]$ and σ_{i1}, σ_{i4} are assigned using (5). σ_{13} and σ_{23} are assigned randomly between -150 and $+150$. We used second order centered finite differences, a ghost point for the Robin boundary conditions and a 20×20 uniform grid for each subdomain. We then run 5 and 10 iterations of (3)–(4), and compare the result with the solution on the whole

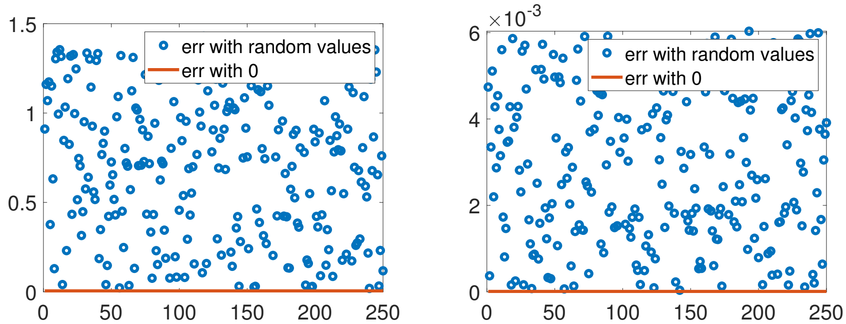

Fig. 1 Number random values between −150 and +150 generated for σ_{i3} VS Error (infinity norm) on λ at the 5th (Left) and 10th (Right) iteration using random antidiagonal elements (blue dots) or just 0 (red line).

domain, which let us compute an error at the end of the iterations. We did this experiment with 250 random couples for σ_{13} and σ_{23}, and plot the error. The results are given in Figure 1. From these results, we see that the choice $\sigma_{13} = \sigma_{23} = 0$ seems to give the lowest error (the red line in (1)). Indeed, after 5 (resp. 10) iterations, the lowest error is at 0.0189 (resp. 3.5919×10^{-5}) with the random values, while the error with $\sigma_{13} = \sigma_{23} = 0$ is at 0.0051 (resp. 8.4667×10^{-6}). This choice is special since it decouples T and λ at the interface $\Gamma_\cap$. It then suggests that the resolution of T first, and then λ, is more efficient, with respect to the number of Schwarz iterations.

3.4 Schwarz iteration as critical points of an optimization problem

We conclude this proceeding by showing that each iteration of the Schwarz method can be obtained by computing the critical points of some specific Lagrange functional. We restrict ourselves to transmission conditions (4), where $\mathcal{S}_i$ are lower-triangular matrices with constant coefficients and recall that $\mathbf{n}$ is the outer normal to $\partial\Omega_1$. We consider the next sub-domain problem

$$\begin{cases} -\mathrm{div}(k\nabla T_i - \mathbf{u}T_i) = f & \text{in } \Omega_i, \\ k\partial_{\mathbf{n}}T_i + a_i\mathbf{u}\cdot\mathbf{n}T_i + (-1)^{i+1}p_iT_i = (-1)^{i+1}g_i & \text{on } \Gamma_\cap,\ T_i = T_0 \text{ on } \partial\Omega_i \setminus \Gamma_\cap. \end{cases} \tag{6}$$

Its variational formulation is: *Find* $T \in H^1(\Omega_i)$ *such that* $T_i|_{\partial\Omega_i\setminus\Gamma_\cap} = T_0$ *and*

$$
\begin{aligned}
a_i(T_i,\lambda_i) :=& \int_{\Omega_i} k\nabla T_i \cdot \nabla\lambda_i - T_i\,\mathbf{u}\cdot\nabla\lambda_i\,dx \\
&- \int_{\Gamma_\cap} \left(p_i + (-1)^{i+1}(1+a_i)\mathbf{u}\cdot\mathbf{n}\right) T_i\lambda_i\,ds \\
&= \int_{\Omega_i} f\,\lambda_i\,dx + \int_{\Gamma_\cap} g_i\lambda_i\,ds,\ \forall\lambda_i \in V_i := \left\{\varphi \in H^1(\Omega_i) \mid \varphi|_{\partial\Omega_i\setminus\Gamma_\cap} = 0\right\}.
\end{aligned}
$$

We then have the next result whose proof can formally be done by direct computation.

Theorem 2 *Given $\alpha_i \in L^\infty(\Gamma_\cap)$, $\beta_i \in L^2(\Gamma_\cap)$, we consider the Lagrangian*

$$
\begin{aligned}
\mathcal{L}_i(T_i,\lambda_i) =& \frac{1}{2}\|T_i - T_{target}\|^2_{L^2(\Omega_i)} + a_i(T_i,\lambda_i) - \int_{\Omega_i} f\,\lambda_i\,dx - \int_{\Gamma_\cap} g_i\lambda_i\,ds \\
&+ \int_{\Gamma_\cap}\left(\frac{\alpha_i}{2}T_i^2 + \beta_i T_i\right)ds,\ \forall\left(T_i - \widetilde{T_{0,i}}\right), \lambda_i \in V_i,
\end{aligned}
$$

where $\widetilde{T_{0,i}} \in V_i$ is an extension of T_0. Let (T_i,λ_i) satisfying $\partial_{T_i,\lambda_i}\mathcal{L}_i(T_i,\lambda_i) = 0$. Then T_i is a weak solution to (6) *and $\lambda_i \in V_i$ is a weak solution to the (adjoint) problem*

$$
\begin{cases}
\operatorname{div}(k\nabla\lambda_i + \mathbf{u}\lambda_i) = T_i - T_{target}\ \text{in}\ \Omega_i,\ \lambda_i = 0\ \text{on}\ \partial\Omega_i\setminus\Gamma_\cap, \\
k\partial_{\mathbf{n}}\lambda_i + (1+a_i)\mathbf{u}\cdot\mathbf{n}\lambda_i + (-1)^{i+1}p_i\lambda_i = (-1)^{i+1}(\alpha_i T_i + \beta_i)\ \text{on}\ \Gamma_\cap.
\end{cases} \tag{7}
$$

From Theorem 2, we see that chosing $a_i = -\frac{1}{2}$, $p_i = \sigma_{i1}$, $\alpha_i = -\sigma_{i3}$,

$$
\begin{aligned}
g_i =& k\partial_{\mathbf{n}}T^n_{3-i} + a_i\mathbf{u}\cdot\mathbf{n}T^n_{3-i} + (-1)^{i+1}p_iT^n_{3-i}, \\
\beta_i =& \left(k\partial_{\mathbf{n}}\lambda^n_{3-i} + (1+a_i)\mathbf{u}\cdot\mathbf{n}\lambda^n_{3-i} + (-1)^{i+1}p_i\lambda^n_{3-i} + \alpha_iT^n_{3-i}\right),
\end{aligned}
$$

yields that $(T_i^{k+1} - \widetilde{T_{0,i}}, \lambda_i^{k+1}) \in V_i \times V_i$ is a critical point of $\mathcal{L}_i$. Each iterate of the DDM can then be obtained by solving an optimization problem on each subdomain (see also e.g. [1, 5]).

4 Conclusion

Using a Schwarz method on (2) appears to be harder than expected. The transmission conditions found in [3] can be adapted to this case, but only gives a partial clue to define some optimized transmission operator. Furthermore, solving (2) only let us compute the gradient of the cost which only accelerates the computation of the gradient, but not necessarily the resolution of (1). Concerning (2), we still wonder if one can take advantage of the triangular structure of (1): is it better to solve first for T alone, and then for λ, or could we find efficient iterations to compute the couple (T,λ)? Additional works in this direction are on-going projects.

Acknowledgements All the authors are supported by the “Agence Nationale de la Recherche” (ANR), Project O-TO-TT-FU number ANR-19-CE40-0011.

References

1. Ciaramella, G., Halpern, L., and Mechelli, L. Convergence analysis and optimization of a Robin Schwarz waveform relaxation method for periodic parabolic optimal control problems. MOX-Report No. 62/2022 (2022).
2. Ciaramella, G., Kwok, F., and Müller, G. Nonlinear optimized Schwarz preconditioner for elliptic optimal control problems. In: *Domain Decomposition Methods in Science and Engineering XXVI*. Springer (2020).
3. Dubois, O. *Optimized Schwarz Methods for the Advection-Diffusion Equation and for Problems with Discontinuous Coefficients*. Ph.D. thesis, McGill University (2007).
4. Gander, M. J. Optimized Schwarz methods. *SIAM Journal on Numerical Analysis* **44**(2), 699–731 (2006).
5. Gong, W., Kwok, F., and Tan, Z. Convergence analysis of the Schwarz alternating method for unconstrained elliptic optimal control problems. *arXiv preprint arXiv:2201.00974* (2022).
6. Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S. *Optimization with PDE constraints*, vol. 23. Springer Science & Business Media (2008).
7. Lagnese, J. E., Leugering, G., and Leugering, G. *Domain in Decomposition Methods in Optimal Control of Partial Differential Equations*. 148. Springer Science & Business Media (2004).
8. Quarteroni, A. and Valli, A. *Domain decomposition methods for partial differential equations*. Oxford University Press (1999).
9. Vieira, A., Bastide, A., and Cocquet, P.-H. Topology optimization for steady-state anisothermal flow targeting solids with piecewise constant thermal diffusivity. *Applied Mathematics & Optimization* **85**(3), 1–32 (2022).