A FETI-INSPIRED APPROACH FOR A LINEAR QUADRATIC OPTIMAL CONTROL PROBLEM.

ALEXANDRE VIEIRA, PIERRE-HENRI COCQUET, MARTIN GANDER

Abstract. This paper introduce a parallel approach to solve a linear quadratic optimal control problem. Instead of parallellizing the resolution of the necessary and sufficient optimality condition, we focus here on a domain decomposition technique appearing directly in the optimization problem. This approach then gives rise to an augmented Lagrangian method in order to deal with the continuity constraints. We prove that, when the algorithm converges, it necessarily converges to an optimal point of the original, non-decomposed problem. Some numerical examples prove the efficiency of this approach.

1. Introduction. The goal of this paper is to find a new parallel approach to solve the following optimal control problem:

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|y - y_{\text{target}}\|^2_{L^2(\Omega)} + \tfrac{\alpha}{2}\|f\|^2_{L^2(\Omega)} \\ s.t. \quad & \begin{cases} -\Delta y = F + f \text{ on } \Omega, \\ y|_{\partial\Omega} = 0, \end{cases} \end{aligned} \tag{1.1}$$

where Ω is a bounded open subset of $\mathbb{R}^d$ for some integer $d \geq 1$, $f \in L^2(\Omega)$ is our optimization parameter, $F \in L^2(\Omega)$ an imposed source term

In order to solve (1.1), we want to split the problem into subproblems involving subdomains only. This could be done by applying a domain decomposition method to solve the necessary and sufficient first-order optimality conditions (see e.g. [19]) which read:

$$\begin{cases} -\Delta p^* = y - y_{\text{target}} \text{ on } \Omega, \\ p^*|_{\partial\Omega} = 0, \\ \alpha f - p^* = 0, \end{cases} \tag{1.2}$$

We emphasize that this approach consists in splitting the computation of the gradient of the reduced cost function on each subdomain ; this approach is efficient only because the control can be computed as a function of the state y^* and the adjoint state p^*. For more difficult problems, involving for instance a non-linear model (see e.g. topology optimization problems [1, 36, 37]), we find this approach inefficient for the following reasons:

(i) from a parallelization point of view, the gradient is computed on both subdomains, but the gradient descent is computed on the whole domain,
(ii) from the efficiency of the iterations: the splitting needs to converge for a fix control f before updating it. This could be relaxed, but it is not yet entirely clear how it is possible.

Our goal in this paper is to split the problem directly at the optimization problem level. First of all, some papers suggest that some decomposition techniques can be seen as optimal control problems. The authors in [14] present the DDM techniques as an optimal control problem, where the controls act as boundary conditions on the subdomains. In [33], the authors express the FETI method, a decomposition technique, as a minimization problem solved through a Lagrangian method. In this article, the analysis of the multiplier associated to the continuity constraints is carefully handled. Other examples of such approach can be found in [3, 12, 20, 23].

Regarding the splitting of the optimization problem, another approach is presented in [19] which splits 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. The generalization of this approach to more difficult problems (with non-linear models, for instance) seems however difficult.

Lagrangian approaches are already used for optimal control problems. In [26], an optimal problem is tackled through a Lagrangian method. The constraint taken into account is the underlying PDE, which is solved iteratively through a penalization in the cost. However, we underline that no domain decomposition is considered. The same approach has been used in [2, 13]. It can be noted that the augmented Lagrangian approach is often used in the literature in order to take into account some constraints that are hardly taken into account in practice [4, 6, 5, 11, 21, 24, 25].

This idea is pushed a bit further in [30], where the authors use a Lagrangian method to take the PDE into the cost, but also use a decomposition technique. However, no proof for convergence is carried in this article. All these articles are of course closely related to the analysis of optimization problem in infinite dimensional spaces ; see for instance [24, 27, 28, 29].

Our approach consists in redefining first the PDE constraint with an equivalent equation defined on several subdomain. This equivalent definition adds continuity constraints that are then tackled with an augmented Lagrangian method. Contrary to [30], we do not augment the cost with the decomposed PDE, which is still considered as an equation defining the state variable. Even if this approach seems really close to the results presented in [9], two major differences can be found:

- we consider the value on the virtual boundary as a control, while they consider it as a multiplier,
- we prove convergence of the proposed algorithm to the optimal solution.

The rest of this article is organized as follows: in section 2, we briefly analyze the underlying PDE system. Section 3 is devoted to the analysis of the Lagrangian algorithm we define. In subsection 4.3, we adopt a Fourier analysis of the optimization problem to measure the rate of convergence of the proposed algorithm. **??** focuses on the discrete counterpart of our problem, which is then illustrated with some numerical examples in section 5.

Notations. We will denote by ∇q the gradient of a real-valued function. Assuming that we have a Hilbert space $\mathcal{H}$ and a subset $\mathcal{X}$ such that $\mathcal{X} \subset \mathcal{H}$. The directional derivative of a function $F : x \in \mathcal{X} \mapsto F(x) \in \mathbb{R}$ is then denoted by

$$\partial_x F(x)[\delta x] = \lim_{t \to 0} \frac{F(x + t\delta x) - F(x)}{t} = \langle \partial_x F(x), \delta x \rangle .$$

where $\partial_x F(x)$ is the gradient of F.
The notation $A \lesssim B$ stands when there exists a constant $C(\Omega)$, which can depend only the domain Ω, such that $A \leq C(\Omega) B$.
For $\Gamma_\cap \subset \partial\Omega$, we denote by $H^s(\Gamma_\cap)$ is the set of restriction to $\Gamma_\cap$ of distributions in $H^s(\partial\Omega)$, and by $H^{\frac{1}{2}}_{00}(\Gamma_\cap)$ the set of distributions defined on $\Gamma_\cap$ such that their extension with 0 on $\partial\Omega$ is in $H^{\frac{3}{2}}(\partial\Omega)$. We refer to [32] for more information about these trace spaces. We will define two subdomains Ω_1 and Ω_2 of a domain Ω such that $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$, $\Omega_1 \cap \Omega_2 = \emptyset$, $\Gamma_\cap = \partial\Omega_1 \cap \partial\Omega_2$. Define $\mathbf{n}_i$ the exterior normal vector to Ω_i. Note that $\mathbf{n}_2 = -\mathbf{n}_1$ on $\Gamma_\cap$.

Based in these subdomains, we will need the functional spaces

$$\mathcal{V}_\cap = H^{\frac{1}{2}}_{00}(\Gamma_\cap),\ \mathcal{U} = L^2(\Omega_1) \times L^2(\Omega_2) \times H^{-\frac{1}{2}}(\Gamma_\cap).$$

2. Analysis of the optimization problem. This section will let us recall some results concerning (1.1). Problem (1.1) is a linear-quadratic optimal control problem which has been analyzed in depth [31]. In this section, we will mainly recall the result concerning the equivalence between a subdomain decomposed problem and the whole domain PDE

$$\begin{cases} -\Delta y = F + f \quad \text{in } \Omega, \\ y|_{\partial\Omega} = 0. \end{cases} \tag{2.1}$$

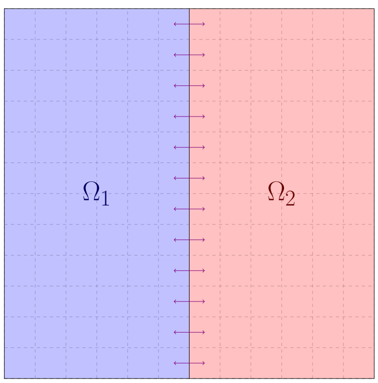


Fig. 2.1: Sketch of the decomposition of a domain Ω in two subdomains.

We restrict the analysis to a decomposition with two subdomains. Denote Ω_1, Ω_2 a splitting such that $\overline{\Omega} = \overline{\Omega_1 \cup \Omega_2}$, $\Omega_1 \cap \Omega_2 = \emptyset$, $\Gamma_\cap = \overline{\Omega_1} \cap \overline{\Omega_2} \neq \emptyset$. We suppose also that $\partial\Omega_i$ has Lipschitz continuity. Consider the equation

$$\begin{cases} -\Delta y_i = f_i & \text{in } \Omega_i, \\ y_i|_{\partial\Omega\cap\partial\Omega_i} = 0, & \\ \partial_{\mathbf{n}} y_i|_{\Gamma_\cap} = (-1)^{i+1} g, & i = 1, 2, \end{cases} \tag{2.2}$$

where we define $\partial_{\mathbf{n}} y_i = \nabla y_i \cdot \mathbf{n}_i$ with $\mathbf{n}_i$ being the outward normal unit vector to Ω_i. We can easily prove that for all $g \in H^{-\frac{1}{2}}(\Gamma_\cap)$, (2.2) has a unique weak solution.

We now state the equivalence existing between (2.1) and (2.2). Denote $\mathcal{V}_i(g) = \{y \in H^1(\Omega_i) | y|_{\partial\Omega_u \setminus \Gamma_\cap} = 0,\ \partial_{\mathbf{n}} y|_{\Gamma_\cap} = (-1)^{i+1} g\}$.

Proposition 2.1. *[35] Let $y \in H^1_0(\Omega)$ be a solution of (2.1). Then for $i = 1, 2$, $y_i = y|_{\Omega_i} \in \mathcal{V}_i(g)$ is a solution of (2.2) with $g = (-1)^{i+1} \partial_{\mathbf{n}} y|_{\Gamma_\cap}$.*
Conversely, define $g \in H^{-\frac{1}{2}}(\Gamma_\cap)$ such that $\mathcal{V}_i(g)$ is not empty for $i = 1, 2$. Let $y_i \in \mathcal{V}_i(g)$ be a solution of (2.2), for $i = 1, 2$. Suppose that $y_1 = y_2$ on $\Gamma_\cap$. Then the function y defined as $y|_{\Omega_i} = y_i$ belongs to $H^1_0(\Omega)$ and is the solution to (2.1).

Remark 2.2. *We emphasize that all the results of this paper involving only the PDE at the continuous level can be generalized to more subdomains as soon as the local boundary conditions ensure the well-posedness of each sub-problems.*

3. A no-overlap decomposed equivalent formulation. In this section, we introduce our domain decomposition of the problem (1.1). We will show that the decomposition introduced in (2.2) adds a continuity constraint that can be relaxed using a Lagrangian method. Our results will focus on the definition of the multiplier associated to the continuity constraints, the analysis of the derivative of the constraints with respect to the controls, and the convergence of the augmented Lagrangian technique. This approach is close to the FETI method, as exposed in [33], where the authors also express the decomposition method as a split optimization problem with a continuity constraint, and where the multiplier associated to the continuity constraint is analyzed. For ease of presentation, we restricted the analysis to the Laplacian operator, but it can be generalized to the more general elliptic operator $\mathcal{A}$, as long as the Theorem 3.2 remains valid for this operator. This is verified if, for example, Weyl's lemma [38, p. 78, Theorem 18.G] still holds for this operator $\mathcal{A}$.

3.1. Problem statement, Lagrangian. We would like now to decompose (1.1) without any overlap. Due to Proposition 2.1, (1.1) is equivalent to:

$$
\begin{aligned}
&\min\ \hat{J}(f,g) = \frac{1}{2}\sum_{i=1}^{2}\|y_i - y_{\text{target}}\|^2_{L^2(\Omega_i)} + \alpha\|f_i\|^2_{L^2(\Omega_i)} \\
&\text{s.t.} \begin{cases} -\Delta y_i = F + f_i & \text{in } \Omega_i, \\ y_i|_{\partial\Omega} = 0, & \\ \partial_{\mathbf{n}} y_i|_{\Gamma_\cap} = (-1)^{i+1} g, & i = 1,2, \\ y_1|_{\Gamma_\cap} = y_2|_{\Gamma_\cap}, & \\ (f_1, f_2, g) \in \mathcal{U}, & \end{cases}
\end{aligned} \tag{3.1}
$$

and (3.1) admits a unique solution since (1.1) does so. We restrict the admissible controls to $\mathcal{U} = L^2(\Omega_1) \times L^2(\Omega_2) \times H^{-\frac{1}{2}}(\Gamma_\cap)$.

We denote the reduced functional $\hat{J}(f,g) = J(y(f,g), f, g)$, where $y(f,g) = (y_1(f_1,g), y_2(f_2,g))$ and $y_i(f_i,g)$ is the solution of (2.2) on Ω_i, $i = 1,2$. It is worth noting that we could have chosen two different $g_i \in H^{-\frac{1}{2}}(\Gamma_\cap)$ in (3.1), with the further constraint $g_1 + g_2 = 0$. The subsequent analysis can be easily adapted to this case.

We need to find the multiplier associated to the continuity of the solution on the interface $\Gamma_\cap$. Note that $y_1|_{\Gamma_\cap} - y_2|_{\Gamma_\cap}$ belongs to the set $\mathcal{V}_\cap = H^{\frac{1}{2}}_{00}(\Gamma_\cap)$. Define $\lambda \in \mathcal{V}^*_\cap$ the multiplier associated to the constraint $y_1 - y_2 = 0$ on $\Gamma_\cap$. The lagrangian of (3.1) reads as:

$$
\begin{aligned}
\mathcal{L}(y,f,g,p,\lambda) = &\sum_{i=1}^{2} \int_{\Omega_i} \nabla y_i \cdot \nabla p_i - (-1)^{i+1}\int_{\Gamma_\cap} g p_i - \int_{\Omega_i} f_i p_i + \frac{1}{2}\int_{\Omega_i} (y_i - y_{\text{target}})^2 \\
&+ \frac{\alpha}{2}\int_{\Omega_i} f_i^2 + \int_{\Gamma_\cap} (y_1 - y_2)\lambda.
\end{aligned}
$$

where we set $f = (f_1, f_2)$, $y = (y_1, y_2)$ and $p = (p_1, p_2)$.

In order to define the adjoint state p_i, we start by differentiating the Lagrangian with respect to y_i and set it to 0, which gives

$$
\partial_{y_i}\mathcal{L}(y,f,g,p,\lambda)[\delta y_i] = \int_{\Omega_i} \nabla\delta y_i \cdot \nabla p_i + \int_{\Omega_i} (y_i - y_{\text{target}})\delta y_i + (-1)^{i+1}\int_{\Gamma_\cap} \delta y_i \lambda = 0.
$$

This is the weak formulation of:

$$
\begin{cases} -\Delta p_i + y_i = y_{\text{target}}, \\ p_i|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}} p_i = (-1)^{i+1}\lambda \text{ on } \Gamma_\cap. \end{cases} \tag{3.2}
$$

Differentiating now with respect to the controls (f,g), one gets:

$$
\partial_{f_i,g}\mathcal{L}(y,f,g,p,\lambda)[\delta f_i, \delta g] = -\int_{\Gamma_\cap} \delta g(p_2 - p_1) - \int_{\Omega_i} \delta f_i p_i + \alpha\int_{\Omega_i} f_i \delta f_i.
$$

which gives the necessary conditions of optimality:

$$\begin{cases} \alpha f_i - p_i = 0, \\ p_2|_{\Gamma_\cap} - p_1|_{\Gamma_\cap} = 0, \end{cases}, \; i = 1, 2. \tag{3.3}$$

We see then that $\lambda = \partial_{\mathbf{n}} p^*|_{\Gamma_\cap}$ is a multiplier that respect the optimality conditions, where p^* was defined in (1.2).

3.2. Regularity of the continuity constraint for the first order condition of optimality. In this section, we will focus on proving that the derivative of the continuity constraint given by $c_o : (f_1, f_2, g) \in \mathcal{U} \mapsto y_1(f_1, g)|_{\Gamma_\cap} - y_2(f_2, g)|_{\Gamma_\cap} \in \mathcal{V}_\cap$, where $y_i(f_i, g)$ solves (2.2), is surjective. This condition (called regularity condition or constraint qualification) serves in order to prove that a minimizer of (3.1) respect the first order conditions of optimality (or KKT conditions). Denote M^* the derivative of c_o with respect to (f_1, f_2, g). We can actually derive an explicit expression of the adjoint of the derivative. Thanks to the definition of $\mathcal{L}$, we notice that

$$\partial_{f,g} \mathcal{L}(y(f,g), f, g, p, \lambda) = \begin{pmatrix} \alpha f_1 - p_1 \\ \alpha f_2 - p_2 \\ p_1 - p_2|_{\Gamma_\cap} \end{pmatrix} = \partial_{f,g} \hat{J}(f,g) + M\lambda,$$

where $M : \mathcal{V}_\cap^* \to \mathcal{U}^*$ is the adjoint of M^*. With the same calculations as above, one proves that:

$$\partial_{f,g} \hat{J}(f,g) = \begin{pmatrix} \alpha f_1 - \bar{p}_1 \\ \alpha f_2 - \bar{p}_2 \\ \bar{p}_2|_{\Gamma_\cap} - \bar{p}_1|_{\Gamma_\cap} \end{pmatrix} \text{ where } \begin{cases} -\Delta \bar{p}_i + y_i = y_{\text{target}}, \\ \bar{p}_i|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}} \bar{p}_i = 0 \text{ on } \Gamma_\cap. \end{cases}$$

Therefore, one has:

$$M\lambda = \partial_{f,g} \mathcal{L}(y(f,g), f, g, p, \lambda) - \partial_{f,g} \hat{J}(f,g) = \begin{pmatrix} q_1 \\ q_2 \\ q_1|_{\Gamma_\cap} - q_2|_{\Gamma_\cap} \end{pmatrix}$$

where $q_i = \bar{p}_i - p_i \in H^1(\Omega_i)$ is the weak solution to the equation:

$$\begin{cases} -\Delta q_i = 0, \\ q_i|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}} q_i = (-1)^{i+1} \lambda \text{ on } \Gamma_\cap. \end{cases} \tag{3.4}$$

It is worth noting that M does not depend on the controls.

LEMMA 3.1. *The operator $M : \lambda \in \mathcal{V}_\cap^* \mapsto M\lambda \in \mathcal{U}^*$ is injective with closed range.*

Proof. Injectivity: Let λ be such that $M\lambda = 0$. The associated q_i are then harmonic and verify the transmission condition (on $\Gamma_\cap$) $q_1|_{\Gamma_\cap} - q_2|_{\Gamma_\cap} = 0$ and $\partial_{\mathbf{n}_1} q_1 + \partial_{\mathbf{n}_2} q_2 = 0$. As a result, $q_i = q|_{\Omega_i}$ where $q \in H_0^1(\Omega)$ is harmonic. Therefore $q = 0$ and thus $\lambda = \partial_{\mathbf{n}_1} q_1 = 0$.

Closed range: Let $M\lambda_n = (q_{1,n}, q_{2,n}, \varphi_n)^\intercal$ be a sequence of images such that $q_{i,n}$ converges toward some q_i in $L^2(\Omega_i)$ and φ_n converges toward some φ in $H_{00}^{1/2}(\Gamma_\cap)$. To show that M has closed range, we have to prove that there exists $\lambda \in H_{00}^{1/2}(\Gamma_\cap)^*$ such that $M\lambda = (q_1, q_2, \varphi)^\intercal$.

Step 1: Since $q_{i,n}$ is harmonic on Ω_i, we have

$$\forall \psi_i \in \mathcal{C}_c^\infty(\Omega_i): \int_{\Omega_i} q_{i,n} \Delta \psi_i \, dx = 0.$$

Using the L^2 convergence of $q_{i,n}$ toward q_i, we obtain that q_i satisfies

$$\forall \psi_i \in \mathcal{C}_c^\infty(\Omega_i): \int_{\Omega_i} q_i \Delta \psi_i \, dx = 0,$$

and Weyl's Lemma (see e.g. [38, p. 78, Theorem 18.G]) ensures that $q_i \in \mathcal{C}^\infty(\Omega_i)$ and satisfies $\Delta q_i = 0$ pointwise in Ω_i.

Step 2: Let us consider the following spaces

$$H_{i,\Delta} := \left\{ \Phi_i \in H^1(\Omega_i) \mid \Delta \Phi_i \in L^2(\Omega_i),\ \Phi_i|_{\partial\Omega_i \setminus \Gamma_\cap} = 0 \right\}.$$

For any $\varphi \in H_{00}^{1/2}(\Gamma_\cap)$, its extension by zero to $\partial\Omega_i$ is in $H^{1/2}(\partial\Omega_i)$ and the surjectivity of the trace operator gives the existence of some $\mathcal{E}\tilde{\varphi} \in H^1(\Omega_i)$ such that $\mathcal{E}\tilde{\varphi}|_{\partial\Omega_i} = \tilde{\varphi}$ and $\|\mathcal{E}\tilde{\varphi}\|_{H^1(\Omega_i)} \lesssim \|\tilde{\varphi}\|_{H_{00}^{1/2}(\partial\Omega_i)}$. From the Green formula, for all $\Phi_i \in H_{i,\Delta}$,

$$\langle \partial_{\mathbf{n}_i} \Phi_i, \varphi \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)} = \int_{\Omega_i} \nabla \Phi_i \cdot \nabla \mathcal{E}\tilde{\varphi} \, dx + \int_{\Omega_i} \Delta \Phi_i \, \mathcal{E}\tilde{\varphi} \, dx,$$

and thus, we can prove, using Cauchy-Schwarz inequality, that

$$\left| \langle \partial_{\mathbf{n}_i} \Phi_i, \varphi \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)} \right| \lesssim \|\varphi\|_{H_{00}^{1/2}(\Gamma_\cap)} \left(\|\Phi_i\|_{H^1(\Omega_i)} + \|\Delta \Phi_i\|_{L^2(\Omega_i)} \right).$$

As a result, taking the supremum over all φ such that $\|\varphi\|_{H_{00}^{1/2}(\Gamma_\cap)} = 1$, we prove that the linear application

$$\partial_{\mathbf{n}_i} : \Phi_i \in H_{i,\Delta} \mapsto \partial_{\mathbf{n}_i} \Phi_i \in H_{00}^{1/2}(\Gamma_\cap)^*$$

is continuous. In addition, for any harmonic $\Phi_i \in H_{i,\Delta}$, we have the bound

$$\|\partial_{\mathbf{n}_i} \Phi_i\|_{H_{00}^{1/2}(\Gamma_\cap)^*} \lesssim \|\Phi_i\|_{H^1(\Omega_i)} \lesssim \|\nabla \Phi_i\|_{L^2(\Omega_i)},$$

where we used Poincaré inequality to get the last upper bound.

Step 3: Both $q_{i,n}$ satisfy the following weak formulation

$$\forall \psi_i \in V_i: \int_{\Omega_i} \nabla q_{i,n} \cdot \nabla \psi_i \, dx = \langle \partial_{\mathbf{n}_i} q_{i,n}, \psi_i \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)},$$

where $V_i := \left\{ \psi_i \in H^1(\Omega_i) \mid \psi_i|_{\partial\Omega_i \setminus \Gamma_\cap} = 0 \right\}$. Now taking $\psi_i = q_{i,n}$, one gets

$$\begin{aligned}
\sum_i \|\nabla q_{i,n}\|_{L^2(\Omega_i)}^2 &= \langle \partial_{\mathbf{n}_i} q_{1,n}, q_{1,n} \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)} + \langle \partial_{\mathbf{n}_2} q_{2,n}, q_{2,n} \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)} \\
&= \langle \partial_{\mathbf{n}_1} q_{1,n}, q_{1,n} - q_{2,n} \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)} \\
&= \langle \partial_{\mathbf{n}_1} q_{1,n}, \varphi_n \rangle_{H_{00}^{1/2}(\Gamma_\cap)^* \times H_{00}^{1/2}(\Gamma_\cap)}
\end{aligned}$$

where we used that $\partial_{\mathbf{n}_1} q_{1,n}|_{\Gamma_\cap} + \partial_{\mathbf{n}_2} q_{2,n}|_{\Gamma_\cap} = 0$ and that thanks to the definition of M, we have $\varphi_n = q_{1,n} - q_{2,n}$. Using now *Step 2*, we obtain that

$$\sum_i \|\nabla q_{i,n}\|_{L^2(\Omega_i)}^2 \leq \|\partial_{\mathbf{n}_i} q_{1,n}\|_{H_{00}^{1/2}(\Gamma_\cap)^*} \|\varphi_n\|_{H_{00}^{1/2}(\Gamma_\cap)} \lesssim \|\nabla q_{1,n}\|_{L^2(\Omega_1)} \|\varphi_n\|_{H_{00}^{1/2}(\Gamma_\cap)}.$$

Using finally Young's inequality, we obtain that

$$\sum_i \|\nabla q_{i,n}\|^2_{L^2(\Omega_i)} \lesssim \|\varphi_n\|^2_{H^{1/2}_{00}(\Gamma_\cap)}.$$

Since φ_n is assumed to converge toward φ, it is bounded and thus the sequences $(q_{i,n})_n$ are also bounded (uniformly with respect to n) in $H^1(\Omega_i)$. We can then extract subsequences that converges weakly in $H^1(\Omega_i)$ toward q_i. The trace operator being compact, we obtain that $q_i|_{\partial\Omega_i\setminus\Gamma_\cap} = 0$ and $\varphi = q_1|_{\Gamma_\cap} - q_2|_{\Gamma_\cap}$. Since $q_i \in V_i$ is harmonic, it also belongs to $H_{i,\Delta}$ and its normal derivative can be defined as an element of $H^{1/2}_{00}(\Gamma_\cap)^*$.

Step 4: We now identify the limit of the sequence λ_n. Let $\eta \in H^{1/2}_{00}(\Gamma_\cap)$, since its extension by zero (denoted by $\tilde\eta$) over $\partial\Omega_i$ is in $H^{1/2}(\partial\Omega_i)$, we have some $\mathcal{E}_i\tilde\eta \in V_i$ such that $\mathcal{E}_i\tilde\eta|_{\partial\Omega_i} = \tilde\eta$ and $\|\mathcal{E}_i\tilde\eta\|_{H^1(\Omega_i)} \lesssim \|\tilde\eta\|_{H^{1/2}_{00}(\Gamma_\cap)}$. Using that $\lambda_n = (-1)^{i+1}\partial_{\mathbf{n}_i} q_{i,n}$ and the weak formulation satisfied by $q_{i,n}$, we have

$$\begin{aligned}\langle \lambda_n, \eta\rangle_{H^{1/2}_{00}(\Gamma_\cap)^*\times H^{1/2}_{00}(\Gamma_\cap)} &= \frac{1}{2}\langle \partial_{\mathbf{n}_1} q_{1,n}, \mathcal{E}_1\tilde\eta|_{\Gamma_\cap}\rangle_{H^{1/2}_{00}(\Gamma_\cap)^*\times H^{1/2}_{00}(\Gamma_\cap)} \\ &\quad -\frac{1}{2}\langle \partial_{\mathbf{n}_2} q_{2,n}, \mathcal{E}_2\tilde\eta|_{\Gamma_\cap}\rangle_{H^{1/2}_{00}(\Gamma_\cap)^*\times H^{1/2}_{00}(\Gamma_\cap)} \\ &= \frac{1}{2}\int_{\Omega_1} \nabla q_{1,n}\cdot\nabla\mathcal{E}_1\tilde\eta\, dx - \frac{1}{2}\int_{\Omega_2}\nabla q_{2,n}\cdot\nabla\mathcal{E}_2\tilde\eta\, dx.\end{aligned}$$

Now passing to the limit (after extracting a subsequence), we obtain that

$$\begin{aligned}\lim_{n\to+\infty}\langle \lambda_n, \eta\rangle_{H^{1/2}_{00}(\Gamma_\cap)^*\times H^{1/2}_{00}(\Gamma_\cap)} &= \frac{1}{2}\int_{\Omega_1}\nabla q_1\cdot\nabla\mathcal{E}_1\tilde\eta\, dx - \frac{1}{2}\int_{\Omega_2}\nabla q_2\cdot\nabla\mathcal{E}_2\tilde\eta\, dx \\ &= \left\langle \frac{1}{2}\left(\partial_{\mathbf{n}_1} q_1 - \partial_{\mathbf{n}_2} q_2\right), \eta\right\rangle_{H^{1/2}_{00}(\Gamma_\cap)^*\times H^{1/2}_{00}(\Gamma_\cap)}.\end{aligned}$$

We have then proved that $(\lambda_n)_n \subset H^{1/2}_{00}(\Gamma_\cap)^*$ has a subsequence converging toward $\lambda = \frac{1}{2}(\partial_{\mathbf{n}_1} q_1 - \partial_{\mathbf{n}_2} q_2)$. We emphasize that each q_i is unique since it is defined as the L^2-limit of $q_{i,n}$ and then the limit of the subsequence of (λ_n) is also unique. Urysohn's subsequence principle finally prove that the whole sequence λ_n converges toward λ. Since $M\lambda = (q_1, q_2, q_1|_{\Gamma_\cap} - q_2|_{\Gamma_\cap})^t = (q_1, q_2, \varphi)^t$ this proves that the range of M is closed. ∎

Applying now [10, Théorème II.20], this proves the

THEOREM 3.2. *The application $M^* : \mathcal{U} \to \mathcal{V}_\cap$ is a surjective linear application.*

4. A FETI-inspired augmented Lagrangian method.

4.1. Existence of solution, algorithm. Based on the previous results, we can now design an augmented Lagrangian method in order to solve (1.1). We introduce an optimization problem where the continuity of y_i across $\Gamma_\cap$ is handled by penalization. This reads as

$$\tag{4.1} \begin{aligned} &\min\ \hat J^+(f, g; \lambda^k, \rho^k) \\ &\text{s.t.} \begin{cases} -\Delta y_i = F + f_i & \text{in } \Omega_i, \\ y_i|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}} y_i|_{\Gamma_\cap} = (-1)^{i+1} g, \quad i = 1, 2, \\ (f_1, f_2, g) \in \mathcal{U}, \end{cases}\end{aligned}$$

where the cost function is

$$\hat{J}^+(f,g;\lambda^k,\rho^k) = \hat{J}(f,g) + \int_{\Gamma_\cap} (y_1 - y_2)\lambda^k + \frac{\rho^k}{2}\|y_1 - y_2\|^2_{L^2(\Gamma_\cap)}.$$

Regarding the properties of (4.3), we have the following result.

THEOREM 4.1. *For any $\lambda^k \in \mathcal{V}_\cap^*$ and $\rho > 0$, the optimization problem (4.1) admits a unique solution.*

Let $(\overline{f},\overline{g})$ be this unique solution and $(\overline{y_1},\overline{y_2})$ satisfying the associated constraints. Then, for any $\lambda^k \in L^2(\Gamma_\cap)$, we have

$$\|\overline{y_1} - \overline{y_2}\|^2_{L^2(\Gamma_\cap)} \lesssim \frac{1}{\rho} + \frac{1}{\rho^2}\left\|\lambda^k\right\|^2_{L^2(\Gamma_\cap)},$$

where the un-appearing constants do not depend on k nor ρ.

Proof. Remark first that the admissible set $\mathcal{U}$ is convex. Furthermore, the application $f_i, g \mapsto y_i(f_i,g)$ is linear and injective. Therefore, we only need to prove the strict convexity of:

$$\mathcal{J}(y,f) = \frac{1}{2}\sum_{i=1}^{2}\left[\|y_i - y_{\text{target}}\|^2_{L^2(\Omega_i)} + \alpha\|f_i\|^2_{L^2(\Omega_i)}\right] + \int_{\Gamma_\cap}(y_1 - y_2)\lambda^k + \frac{\rho}{2}\|y_1 - y_2\|^2_{L^2(\Gamma_\cap)}.$$

$\mathcal{J}$ is twice differentiable and its Hessian matrix is obviously positive definite. It proves that $\mathcal{J}$ is strictly convex and [22, Theorem 1.46] proves the result.

Let $(\overline{f},\overline{g})$ be the unique solution to (4.1). We then have

$$\mathcal{J}\left(\overline{y}(\overline{f},\overline{g}),\overline{f}\right) \le \mathcal{J}(y(f,g),f), \text{ for all admissible } f,g. \tag{4.2}$$

For any fixed $f \in L^2(\Omega)$, let y_Ω be the unique solution to (2.1) and g_Ω such that the restriction $y_{\Omega,i} := y_\Omega|_{\Omega_i}$ satisfy $\partial_{\mathbf{n}_\mathcal{A}} y_i|_{\Gamma_\cap} = (-1)^{i+1} g_\Omega$. Since $y_\Omega \in H^1(\Omega)$, we have $y_{\Omega,1} = y_{\Omega,2}$ on $\Gamma_\cap$ and (4.2) then yields

$$\begin{aligned}\int_{\Gamma_\cap}(\overline{y_1} - \overline{y_2})\lambda^k + \frac{\rho}{2}\|\overline{y_1} - \overline{y_2}\|^2_{L^2(\Gamma_\cap)} &\le \mathcal{J}(y_\Omega(f,g_\Omega),f)\\ &= \frac{1}{2}\left[\|y_\Omega - y_{\text{target}}\|^2_{L^2(\Omega)} + \alpha\|f\|^2_{L^2(\Omega)}\right] = C.\end{aligned}$$

Young's inequality together with Cauchy-Scwharz inequality then give

$$\begin{aligned}2\left|\int_{\Gamma_\cap}(\overline{y_1} - \overline{y_2})\lambda^k\right| &\le 2\left\|\lambda^k\right\|_{L^2(\Gamma_\cap)}\|\overline{y_1} - \overline{y_2}\|_{L^2(\Gamma_\cap)}\\ &\le \frac{2}{\rho}\left\|\lambda^k\right\|^2_{L^2(\Gamma_\cap)} + \frac{\rho}{2}\|\overline{y_1} - \overline{y_2}\|^2_{L^2(\Gamma_\cap)}\end{aligned}$$

Gathering the previous inequalities, we get

$$\begin{aligned}\rho\|\overline{y_1} - \overline{y_2}\|^2_{L^2(\Gamma_\cap)} &\le 2C - 2\int_{\Gamma_\cap}(\overline{y_1} - \overline{y_2})\lambda^k\\ &\le 2C + \frac{2}{\rho}\left\|\lambda^k\right\|^2_{L^2(\Gamma_\cap)} + \frac{\rho}{2}\|\overline{y_1} - \overline{y_2}\|^2_{L^2(\Gamma_\cap)},\end{aligned}$$

from which we finally infer the aforementioned estimate. ∎

We now compute the stationary conditions of (4.1). The adjoint variables satisfy:

$$
\begin{cases} -\Delta \tilde{p}_i + y_i = y_{\text{target}}, \\ \quad \tilde{p}_i|_{\partial\Omega} = 0, \\ \quad \partial_{\mathbf{n}} \tilde{p}_i = (-1)^i \left(\lambda^k + \rho(y_1 - y_2)\right) \text{ on } \Gamma_\cap, \end{cases} \tag{4.3}
$$

and the necessary and sufficient conditions of optimality for (4.1) then read as:

$$
\begin{cases} \alpha f_i - \tilde{p}_i = 0, \; i = 1, 2, \\ \tilde{p}_2|_{\Gamma_\cap} - \tilde{p}_1|_{\Gamma_\cap} = 0. \end{cases}
$$

We define then the Algorithm 4.1 in order to solve (1.1) by using a domain decomposition at the optimization level. In this algorithm, note that one has:

$$
\partial_{f,g} \hat{J}^+(f^k, g^k; \lambda^k; \rho^k) = \partial_{f,g} \hat{J}(f^k, g^k) + M\bar{\lambda}(f^k, g^k, \lambda^k, \rho^k). \tag{4.4}
$$

We emphasize that $M\bar{\lambda} = M(\lambda^k + \rho^k(y_1(f_1^k, g^k)|_{\Gamma_\cap} - y_2(f_2^k, g^k)|_{\Gamma_\cap}))$ is well defined, since $y_1(f_1^k, g^k)|_{\Gamma_\cap} - y_2(f_2^k, g^k)|_{\Gamma_\cap} \in \mathcal{V}_\cap \hookrightarrow \mathcal{V}_\cap^*$.

Algorithm 4.1: Augmented Lagrangian algorithm

Data: $\rho^0 \geq 1$, $\omega^* << 1$, $\eta_* << 1$, $\tau > 1$.
Choose an initial f^0, g^0, λ^0.
while $\|y_1^k - y_2^k\|^2_{L^2(\Gamma_\cap)} \geq \eta_*$, $\|\partial_{f,g} \hat{J}^+(f^k, g^k)\| \geq \omega_*$ **do**
Solve approximately (4.1) to find f^k, g^k and the associated $\tilde{p}_i^k$, in the sense that :

$$
\|\partial_{f,g} \hat{J}^+(f^k, g^k)\| = \sum_{i=1}^{2} \|\alpha f_i^k - \tilde{p}_i\|^2_{L^2(\Omega_i)} + \|\tilde{p}_1^k - \tilde{p}_2^k\|^2_{L^2(\Gamma_\cap)} \leq \omega_k.
$$

if $\|y_1 - y_2\|^2_{L^2(\Gamma_\cap)} \leq \eta_k$ **then**
`// Update multiplier;`
Choose $\lambda^{k+1} = \bar{\lambda}(f^k, g^k, \lambda^k, \rho^k) = \lambda^k + \rho^k(y_1^k - y_2^k)$;
Let ρ^k unchanged : $\rho^{k+1} = \rho^k$;
Decrease ω_k : $\omega_{k+1} = (\rho^k)^{-1}\omega_k$;
Decrease η_k : $\eta_{k+1} = (\rho^k)^{-1/2}\omega_k$;
else
`// Increase penalization;`
λ^k remains unchanged;
Increase ρ^k: $\rho^{k+1} = \tau\rho^k$;
Decrease ω_k : $\omega_{k+1} = (\rho^{k+1})^{-1}$;
Decrease η_k : $\eta_{k+1} = (\rho^{k+1})^{-1/2}$;
end
end

Remark 4.2. *Note that in Algorithm 4.1, the resolution of (4.1) is parallel. For fixed f_i and g, the resolution of the PDE-constraint can be done in parallel. The same hold for the resolution of the adjoint system (4.3). Also, during the update of* *the descent method, the update of the f_i can be done in parallel, and most of the computation of $\hat{J}^+$ is decoupled on each subdomain. The only non-parallel part is done for the computation of the cost on the virtual boundary $\Gamma_\cap$ and the update of the virtual control g.*

4.2. Convergence of the algorithm. For the analysis of the Algorithm 4.1, we will denote $x = (f_1, f_2, g)$. As in [28] (see also references therein), we will prove first that an asymptotic KKT condition implies convergence to the optimal solution of (3.1). We will prove afterwards that Algorithm 4.1 implies that the generated sequence complies with this asymptotic KKT condition which is defined below.

DEFINITION 4.3. *We say that a feasible point $x \in \mathcal{U}$ respect the AKKT condition if there are sequences $x_k \to x$ in $\mathcal{U}$ and $(\lambda^k) \subset \mathcal{V}_\cap^*$ such that*

$$\partial_x \mathcal{L}(y(x^k), x^k, p(x^k), \lambda^k) = \partial_x \hat{J}(x^k) + M\lambda^k \xrightarrow[k\to+\infty]{} 0 \text{ in } \mathcal{U}^*. \tag{4.5}$$

A first consequence of this convergence property is the boundedness of the sequence of multipliers.

LEMMA 4.4. *Suppose (4.5) is verified at some point x^* (not necessarily admissible), and (x^k, λ^k) be the associated sequence. Then (λ^k) is bounded.*

Proof. First, note that $x \mapsto \partial_x \hat{J}(x)$ and $\lambda \mapsto M\lambda$ are continuous.

Owing to Theorem 3.2 and [10, Théorème II.2.20], the following estimate holds:

$$\forall \lambda \in \mathcal{V}_\cap^*, \ \|\lambda\|_{\mathcal{V}_\cap^*} \lesssim \|M\lambda\|_{\mathcal{U}}^*. \tag{4.6}$$

Therefore, for any $n \in \mathbb{N}$.

$$\begin{aligned}
\|\lambda^n - \lambda^0\|_{\mathcal{V}_\cap^*} &\lesssim \|M\lambda^n - M\lambda^0\|_{\mathcal{U}^*} \\
&\lesssim \|\partial_x \hat{J}(x^n) + M\lambda^n\|_{\mathcal{U}^*} + \|\partial_x \hat{J}(x^0) + M\lambda^0\|_{\mathcal{U}^*} + \|\partial_x \hat{J}(x^n) - \partial_x \hat{J}(x^0)\|_{\mathcal{U}^*} \\
&\leq C.
\end{aligned}$$

∎

We now prove that an AKKT point x^* is the optimal solution.

PROPOSITION 4.5. *Suppose x^* is an AKKT point and denote by (x^k, λ^k) its associated sequence. Then x^* is the optimal solution of (3.1) with associated multiplier λ^* which is the weak-limit of (λ^k).*

Proof. Since $\mathcal{V}_\cap^*$ is reflexive, its unit ball is weakly compact. Therefore, Lemma 4.4 implies that there exists λ^* such that $\lambda^n \rightharpoonup \lambda^*$ in $\mathcal{V}_\cap^*$. Therefore, taking the limit in (4.5), and since x^* is admissible, (x^*, λ^*) is such that

$$\partial_x \mathcal{L}(y(x^*), x^*, p(x^*), \lambda^*) = \partial_x \hat{J}(x^*) + M\lambda^* = 0.$$

Since the problem (3.1) is convex, it implies that x^* is the optimal solution of (3.1).∎

We now just need to prove that Algorithm 4.1 produces a sequence $(x^k, \lambda^k)_k$ which converges to an AKKT point. Combined with Proposition 4.5, this will prove the convergence of the Algorithm 4.1 to a solution of the original problem (1.1). We recall that, in this algorithm, x^{k+1} must be such that:

$$\|\partial_x \hat{J}^+(x^{k+1}; \lambda^k, \rho^k)\|_{\mathcal{U}^*} = \|\partial_{f,g} \hat{J}(x^{k+1}) + M\lambda^{k+1}\|_{\mathcal{U}^*} \leq \omega_k \tag{4.7}$$

As proved in [34, Lemma 3.1.1], $\lim \omega_k = \lim \eta_k = 0$. Remark, with the help of (4.4), that (4.7) implies (4.5) (since we defined $\lambda^{k+1} = \bar{\lambda}$). We therefore only need to prove that the algorithm converges to an admissible point.

Proposition 4.6. *Denote (x^k, λ^k, ρ^k) generated by the Algorithm 4.1, and suppose that $x^k \to x^*$ in $\mathcal{U}$. Then $y_1(x^k) - y_2(x^k) \to 0$, and x^* is an AKKT point.*

Proof. There are two cases : (ρ^k) bounded, and $\rho^k \to +\infty$.

1. Suppose (ρ^k) bounded. Due to how (ρ^k) is updated, it implies that there exists $k_0 \in \mathbb{N}$ such that, for all $k \geq k_0$, $\rho^k = \rho^{k_0}$. Therefore, it implies that for all $k \geq k_0$, $\|y_1(x^k) - y_2(x^k)\|_{L^2(\Gamma_\cap)} \leq \eta_k$. Since $\eta_k \to 0$, this implies $y_1(x^k) - y_2(x^k) \to 0$ in $L^2(\Gamma_\cap)$.
2. Let us now suppose $\rho^k \to +\infty$. Therefore, dividing (4.7) with ρ^k for k large enough,

$$(\rho^k)^{-1}\left(\partial_x \hat{J}(x^{k+1}) + M\lambda^k\right) + M\left(y_1(x^{k+1}) - y_2(x^{k+1})\right) \to 0$$

Since $(\partial_x \hat{J}(x^k))$ and (λ^k) are bounded, we find

$$M\left(y_1(x^{k+1}) - y_2(x^{k+1})\right) \to 0.$$

Using now (4.6), one finds $y_1(x^k) - y_2(x^k) \to 0$ in $\mathcal{V}_\cap^*$.

Now that we have proved that x^* is admissible, (4.7) implies that x^* is an AKKT point. ∎

4.3. Fourier analysis for constant penalization parameter. In this section, we study the rate of convergence of Algorithm 4.1 in the case where the penalization parameter ρ is kept constant during the iterations.

We are going to restrict the optimization problem to the unbounded domain $\Omega = \mathbb{R}^2$ with interface $\Gamma_\cap = \{0\} \times \mathbb{R}$. In such setting, we can study the error of the algorithm, namely how fast the iterations on $\{\lambda^k\}$ are converging to 0, thanks to some Fourier analysis. We refer for instance to [16, 17] for applications of this technique to study the convergence of domain decomposition methods.

Let $\Omega_1 = (-\infty, 0) \times \mathbb{R}$, $\Omega_2 = (0, +\infty) \times \mathbb{R}$ be the two subdomains and assume that λ^0 has enough regularity so that its Fourier transform along the vertical axis is well-defined. The necessary and sufficient optimality conditions for (4.1) read

$$\begin{cases} -\Delta y_i^k = F + \alpha^{-1} p_i^k & \text{in } \Omega_i, \\ y_i^k|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}_i} y_i^k|_{\Gamma_\cap} = (-1)^{i+1} g^k, & i = 1, 2 \end{cases}$$

$$\begin{cases} -\Delta p_i^k + y_i^k = y_{\text{target}}, \\ p_i^k|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}_i} p_i^k = (-1)^i \left(\lambda^k + \rho(y_1^k - y_2^k)\right) \text{ on } \Gamma_\cap, \\ p_1^k = p_2^k \text{ on } \Gamma_\cap. \end{cases}$$

Denote (x_1, x_2) a system of coordinates on $\mathbb{R}^2$. Since we are interested in the error associated to Algorithm 4.1, without loss of generality, we will suppose that $F = 0$ and $y_{\text{target}} = 0$. Let ω the Fourier variable and $(\hat{y}, \hat{p})$ be the Fourier transform in the x_2 direction of (y, p). The above equations then become:

$$\begin{cases} (-\partial_{x_1}^2 + \omega^2)\hat{y}_i^k = \alpha^{-1}\hat{p}_i^k & \text{in } \Omega_i, \\ \hat{y}_i^k|_{\partial\Omega} = 0, \\ \partial_{x_1}\hat{y}_i^k|_{\Gamma_\cap} = \hat{g}^k, & i = 1, 2 \end{cases}$$

$$\begin{cases} (-\partial_{x_1}^2 + \omega^2)\hat{p}_i^k + \hat{y}_i^k = 0, \\ \hat{p}_i^k|_{\partial\Omega} = 0, \\ \partial_{x_1}\hat{p}_i^k = -\left(\hat{\lambda}^k + \rho(\hat{y}_1^k - \hat{y}_2^k)\right) \text{ on } \Gamma_\cap, \end{cases}$$

$$\hat{p}_1^k = \hat{p}_2^k \text{ on } \Gamma_\cap.$$

We compute $\partial_{x_1}^4 \hat{y}_i^k$:

$$\begin{aligned} \partial_{x_1}^4 \hat{y}_i^k =& \omega^2 \partial_{x_1}^2 \hat{y}_i^k - \alpha^{-1}\partial_{x_1}^2 \hat{p}_i^k \\ =& \omega^2 \partial_{x_1}^2 \hat{y}_i^k - \alpha^{-1}(\omega^2 \hat{p}_i^k + \hat{y}_i^k) \\ =& \omega^2 \partial_{x_1}^2 \hat{y}_i^k - \alpha^{-1}\hat{y}_i^k - \omega^2(-\partial_{x_1}^2 \hat{y}_i^k + \omega^2 \hat{y}_i^k) \\ =& 2\omega^2 \partial_{x_1}^2 \hat{y}_i^k - (\omega^4 + \alpha^{-1})\hat{y}_i^k. \end{aligned}$$

Using the boundary conditions at infinity, it admits the solutions

$$\begin{aligned} \hat{y}_1^k =& c_1^k \exp\left(x_1\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) + c_3^k \exp\left(x_1\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right), \\ \hat{y}_2^k =& c_2^k \exp\left(-x_1\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) + c_4^k \exp\left(-x_1\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right). \end{aligned}$$

This implies

$$\begin{aligned} \hat{p}_1^k =& \mathrm{i}\alpha^{\frac{1}{2}}\left(-c_1^k \exp\left(x_1\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) + c_3^k \exp\left(x_1\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right)\right), \\ \hat{p}_2^k =& \mathrm{i}\alpha^{\frac{1}{2}}\left(-c_2^k \exp\left(-x_1\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) + c_4^k \exp\left(-x_1\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right)\right). \end{aligned}$$

The transmission conditions at $x_1 = 0$ then yield

$$\begin{aligned} \partial_{x_1}\hat{y}_1^k|_{x_1=0} =& c_1^k\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} + c_3^k\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} = \hat{g}^k, \\ \partial_{x_1}\hat{y}_2^k|_{x_1=0} =& -c_2^k\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} - c_4^k\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} = \hat{g}^k, \\ \partial_{x_1}\hat{p}_1^k|_{x_1=0} =& -\mathrm{i}\alpha^{\frac{1}{2}}\left(c_1^k\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} - c_3^k\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) = -\hat{\lambda}^{k+1}, \\ \partial_{x_1}\hat{p}_2^k|_{x_1=0} =& \mathrm{i}\alpha^{\frac{1}{2}}\left(c_2^k\sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2} - c_4^k\sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}\right) = -\hat{\lambda}^{k+1}. \end{aligned}$$

Denoting $A = \sqrt{\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}$ and $B = \sqrt{-\mathrm{i}\alpha^{\frac{-1}{2}} + \omega^2}$, We now need to solve the linear system

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} Ac_1^k \\ Ac_2^k \\ Bc_3^k \\ Bc_4^k \end{pmatrix} = \begin{pmatrix} \hat{g}^k \\ \hat{g}^k \\ \mathrm{i}\alpha^{-\frac{1}{2}}\hat{\lambda}^{k+1} \\ -\mathrm{i}\alpha^{-\frac{1}{2}}\hat{\lambda}^{k+1} \end{pmatrix}$$

The inverse of this matrix being simply half its transpose, this proves that:

$$\begin{aligned} Ac_1^k = -Ac_2^k &= \frac{1}{2}(\hat{g}^k + \mathrm{i}\alpha^{-\frac{1}{2}}\hat{\lambda}^{k+1}) \\ Bc_3^k = -Bc_4^k &= \frac{1}{2}(\hat{g}^k - \mathrm{i}\alpha^{-\frac{1}{2}}\hat{\lambda}^{k+1}) \end{aligned}$$

Note in particular that $c_3^k = -c_4^k$ and $c_1^k = -c_2^k$. Furthermore, note that

$$\begin{aligned}
\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0} =& c_1^k + c_3^k - (c_2^k + c_4^k)\\
=& 2(c_1^k + c_3)\\
=& (A^{-1} + B^{-1})\hat{g}^k + (A^{-1} - B^{-1})\mathrm{i}\alpha^{-\frac{1}{2}}\hat{\lambda}^{k+1}.\\
\hat{p}_1^k|_{x_1=0} - \hat{p}_2^k|_{x_1=0} =& -\mathrm{i}\alpha^{\frac{1}{2}}(c_1^k - c_3^k) + \mathrm{i}\alpha^{\frac{1}{2}}(c_2^k - c_4^k)\\
=& -\mathrm{i}\alpha^{\frac{1}{2}}(c_1^k - c_3^k - c_2^k + c_4^k)\\
=& 2\mathrm{i}\alpha^{\frac{1}{2}}(c_3^k - c_1^k)\\
=& \mathrm{i}\alpha^{\frac{1}{2}}(B^{-1} - A^{-1})\hat{g}^k + (A^{-1} + B^{-1})\hat{\lambda}^{k+1}.
\end{aligned}$$

Since $\hat{p}_1^k|_{x_1=0} - \hat{p}_2^k|_{x_1=0} = 0$, this implies $\hat{g}^k = \mathrm{i}\alpha^{-\frac{1}{2}}\frac{(A+B)}{A-B}\hat{\lambda}^{k+1}$ from which we infer

$$\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0} = 4\mathrm{i}\alpha^{-\frac{1}{2}}(A-B)^{-1}\hat{\lambda}^{k+1}.$$

Denote $D = 4\mathrm{i}\alpha^{-\frac{1}{2}}(A-B)^{-1}$. On top of that, we define $\hat{\lambda}^{k+1}$ with the recurrence $\hat{\lambda}^k = \hat{\lambda}^{k+1} - \rho(\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0})$. Therefore, substituting the expression of $\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0}$ gives $\hat{\lambda}^k = (1-\rho D)\hat{\lambda}^{k+1}$, and we finally obtain the following convergence factor $R := \frac{\hat{\lambda}^{k+1}}{\hat{\lambda}^k} = (1-\rho D)^{-1}$. Using the expressions for D, we have

$$R = \left(1 - \frac{4\mathrm{i}\rho}{\alpha^{1/2}\left(\sqrt{\omega^2 + \mathrm{i}\alpha^{-1/2}} - \sqrt{\omega^2 - \mathrm{i}\alpha^{-1/2}}\right)}\right)^{-1}.$$

Some computations then give

$$\lim_{\alpha\to 0} R = 0,\ R = \frac{\mathrm{i}\alpha^{1/2}}{4\rho}\left(\sqrt{\omega^2 + \mathrm{i}\alpha^{-1/2}} - \sqrt{\omega^2 - \mathrm{i}\alpha^{-1/2}}\right) + O\left(\frac{1}{\rho^2}\right).$$

We emphasize that, for all $\alpha \geq 0$ and any frequency ω, $|R|$ is strictly smaller than 1 as soon as ρ large enough.

Also we have:

$$\frac{\hat{\lambda}^{k+1} - \hat{\lambda}^k}{\hat{\lambda}^k - \hat{\lambda}^{k-1}} = \frac{\rho(\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0})}{\rho(\hat{y}_1^{k-1}|_{x_1=0} - \hat{y}_2^{k-1}|_{x_1=0})} = \frac{\rho D\hat{\lambda}^{k+1}}{\rho D\hat{\lambda}^k} = R.$$

This shows that the rate of convergence on $\{\lambda^k\}$ is directly linked to the rate of convergence of $\{\hat{y}_1^k|_{x_1=0} - \hat{y}_2^k|_{x_1=0}\}_k$. Therefore, the smaller $|R|$ is, the faster the convergence of $\{y_1^k|_{x_1=0} - y_2^k|_{x_1=0}\}_k$ (or similarly on $\{\lambda^k\}$) is. We show a plot of $|R|$ for chosen parameters in Figure 4.1.

4.4. Finite element method. We now present the discrete counterpart of the Lagrangian method introduced in section 3. We use a standard Galerkin method in order to transform the continuous problem (3.1) to the following problem (using the notations of [15])

$$\begin{aligned}
&\min \sum_{i=1}^{2} \frac{1}{2}(u_i - u_{\text{target}_i})^\intercal N_i(u_i - u_{\text{target}_i}) + \frac{\alpha}{2} f_i^\intercal N_i f_i\\
&\text{s.t.} \begin{cases} K_i u_i = F_i + N_i f_i + (-1)^{i+1} B_i^\intercal g,\ i = 1,2\\ B_1 u_1 - B_2 u_2 = 0, \end{cases}
\end{aligned} \tag{4.8}$$

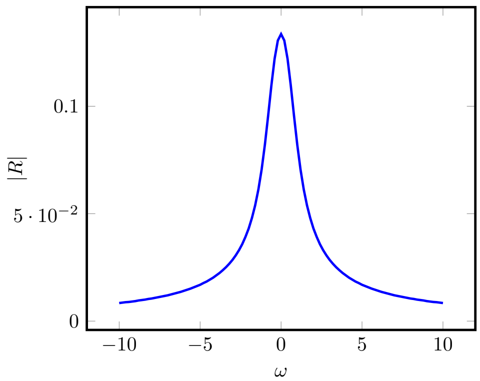


Fig. 4.1: $|R|$ for $\rho = 3$, $\alpha = 1$.

where $u_i \in \mathbb{R}^{n_i^s + n^I}$, where n_i^s is the number of interior node in Ω_i and n^I the number of interface node on $\Gamma_\cap$, and B_i is of the form $B_i = [0_i, I_i]$, $i = 1, 2$, where 0_i is an $n^I \times n_i^s$ null matrix and I_i is the $n^I \times n^I$ identity matrix. K_i is the stiffness matrix and N_i the interpolation matrix, that we suppose to be both invertible.

We solve (4.8) with an augmented lagrangian method:

$$
\begin{aligned}
\min\ & \sum_{i=1}^{2} \frac{1}{2}(u_i - u_{\text{target}_i})^\intercal N_i (u_i - u_{\text{target}_i}) + \frac{\alpha}{2} f_i^\intercal N_i f_i + (\lambda^k)^\intercal (B_1 u_1 - B_2 u_2) \\
& + \frac{\rho^k}{2} \|B_1 u_1 - B_2 u_2\|_2^2 \\
\text{s.t. } & K_i u_i = F_i + N_i f_i + (-1)^{i+1} B_i^\intercal g,\ i = 1, 2
\end{aligned} \tag{4.9}
$$

As in the continuous framework, the convergence of (4.9) depends on the surjectivity of the derivative of $M_h : (f_1, f_2, g) \mapsto B_1 u_1 - B_2 u_2$.

Proposition 4.7. *The derivative of M_h w.r.t. (f_1, f_2, g) is onto.*

Proof. Note first that

$$
B_1 u_1 - B_2 u_2 = \begin{pmatrix} B_1 K_1^{-1} N_1 & -B_2 K_2^{-1} N_2 & B_1 K_1^{-1} B_1^\intercal - B_2 K_2^{-1} B_2^\intercal \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ g \end{pmatrix}
$$
$$
+ B_1 K_1^{-1} F_1 - B_2 K_2^{-1} F_2.
$$

Therefore, the derivative of M_h is simply

$$
\partial M_h = \begin{pmatrix} B_1 K_1^{-1} N_1 & -B_2 K_2^{-1} N_2 & B_1 K_1^{-1} B_1^\intercal - B_2 K_2^{-1} B_2^\intercal \end{pmatrix}.
$$

In order to prove that ∂M_h is surjective, we will prove that $(\partial M_h)^\intercal$ is injective. Let $\lambda \in \mathbb{R}^{n^I}$. Suppose $(\partial M_h)^\intercal \lambda = 0$. This implies:

$$
\begin{aligned}
& K_1^{-\intercal} B_1^\intercal N_1 \lambda = 0 \\
& - K_2^{-\intercal} B_2^\intercal N_2 \lambda = 0 \\
& (B_1 K_1^{-\intercal} B_1^\intercal - B_2 K_2^{-\intercal} B_2^\intercal)\lambda = 0
\end{aligned}
\iff
\begin{aligned}
& B_1^\intercal N_1 \lambda = 0 \\
& B_2^\intercal N_2 \lambda = 0
\end{aligned}
\iff \lambda = 0
$$
∎

The proof of the convergence of this discrete lagrangian approach (4.9) to the solutions of (4.8) can then be done as in the previous section or can be found in [7, 8].

5. Numerical example. We now show how the above results work on a simple example, inspired by [19]. Namely, we want to solve numerically the following problem:

$$\begin{aligned} &\min \frac{1}{2}\|y - y_{\text{target}}\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|f\|^2_{L^2(\Omega)} \\ &\text{s.t.} \begin{cases} -\Delta y = F + f \text{ on } \Omega, \\ y|_{\partial\Omega} = 0, \end{cases} \end{aligned} \tag{5.1}$$

where $\Omega = [-1, 1] \times [0, 1]$, $y_{\text{target}}(x_1, x_2) = C \sin(k\pi x_1) \sin(k\pi x_2)$ and $F(x_1, x_2) = 2Ck^2\pi^2 \sin(k\pi x_1) \sin(k\pi x_2)$, for different parameters $C \in \mathbb{R}$, $\alpha > 0$, $k \in \mathbb{N}$. The optimal solution is $f^* = 0$, $y^* = y_{\text{target}}$ for all admissible choices of parameters C, k, α. This problem is solved with Algorithm 4.1 using $\eta^* = \omega^* = 10^{-10}$.

The Poisson equation is discretized using a finite element scheme. The virtual border $\Gamma_\cap$ is $\{0\} \times [0, 1]$. The finite element discretisation is built using Q1 elements on a structured mesh with constant step size h in both directions x_1 and x_2, chosen such that $\Gamma_\cap$ is composed of facets (and not included inside cells). The minimization problems are then solved using the *quadprog* routine of MATLAB, which uses an interior-point algorithm. The code used for the computations is available online[1].

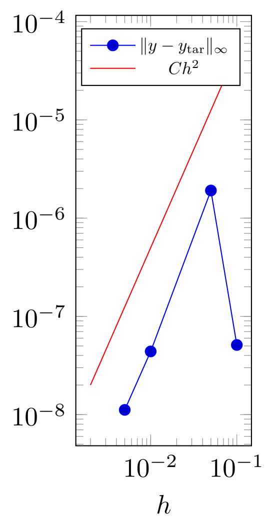
First, we test how the algorithm behaves with different values of α. This is summarised in Table 5.1. Overall, the algorithm remains stable with the different values of α. We notice that the algorithm converges in a few number of iterations.

α	$\|y - y_{\text{target}}\|_\infty$	$\|f\|_\infty$	$\|g - g_{\text{target}}\|_\infty$	Cost	# iterations
10	$1.41 \cdot 10^{-7}$	$1.47 \cdot 10^{-8}$	$3.69 \cdot 10^{-7}$	$1.98 \cdot 10^{-13}$	3
1	$4.4 \cdot 10^{-8}$	$1.91 \cdot 10^{-8}$	$3.99 \cdot 10^{-7}$	$1.93 \cdot 10^{-14}$	3
$1 \cdot 10^{-2}$	$9.1 \cdot 10^{-9}$	$1.39 \cdot 10^{-8}$	$4.17 \cdot 10^{-7}$	$8.29 \cdot 10^{-16}$	3
$1 \cdot 10^{-4}$	$8.33 \cdot 10^{-9}$	$1.33 \cdot 10^{-8}$	$4.17 \cdot 10^{-7}$	$6.96 \cdot 10^{-16}$	3
$1 \cdot 10^{-6}$	$8.33 \cdot 10^{-9}$	$1.33 \cdot 10^{-8}$	$4.17 \cdot 10^{-7}$	$6.94 \cdot 10^{-16}$	3

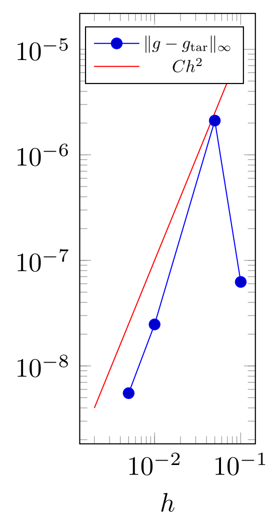
Table 5.1: Error on the solution using different values of α and a finite element scheme. $k = 2$, $C = 1$, $h = 1.10^{-2}$.

In Figure 5.1, we plot the error of different variables with respect to the discretization step. As it can be expected, since we used a second order discretization of the equation (2.2) (see also [18]), we remark that the error decrease is of order 2. We remark also that the number of iterations remains stable when the step size becomes smaller, even though it implies a growing number of decision variables.

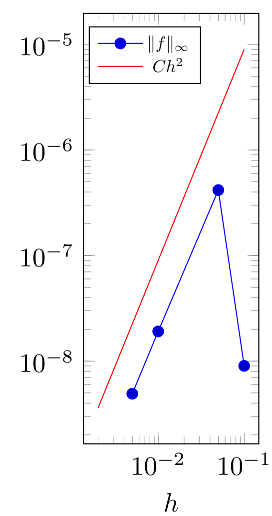
[1] https://gitlab.osureunion.fr/avieira/liquofeti



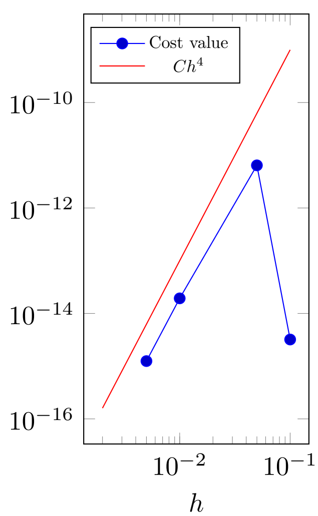
(a) Error on the state function.



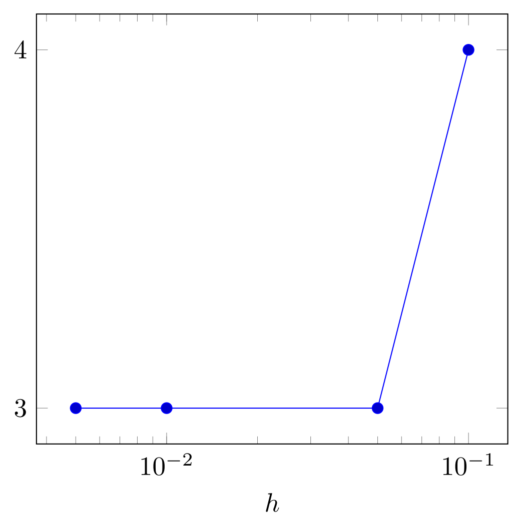
(b) Error on the normal derivative at the interface.



(c) Error on the control function.



(d) Value of the cost.



(e) Number of iterations before convergence.

Fig. 5.1: Errors using different discretization steps ($\alpha = 1$, $C = 1$).

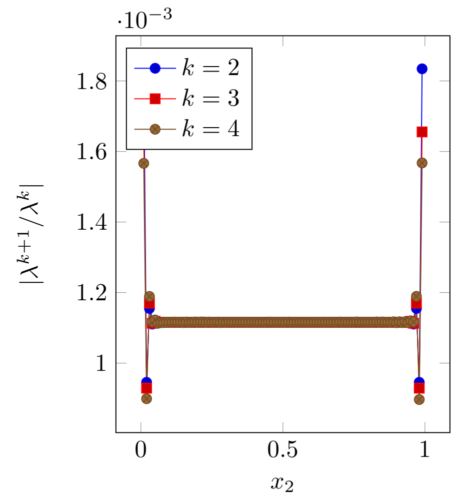


Fig. 5.2: $|\lambda^{k+1}/\lambda^k|$ for different iterations, using finite element.

Finally, we measure the rate of convergence on λ in order to compare the experimental results with the analysis of subsection 4.3. We choose $C = 0$, $k = 1$, $\alpha = 1$, and start the iterations with $\lambda^0 = \sin(\pi x_2)$, $\rho = 3$. With these parameters, the analysis of subsection 4.3 predicts a converging factor of approximately $2.5.10^{-2}$. As shown in Figure 5.2, we get a converging factor even smaller than predicted, around 10^{-3}, showing the fast convergence of the algorithm. Different tests using a finite difference scheme are presented in Appendix A.2.

6. Conclusions. We have designed a parallel approach to solve a linear quadratic optimal control problem involving the Laplace operator. This approach, inspired by the FETI and the Lagrangian methods for finding constrained optima, let us have a powerful approach to solve optimal control problems in a parallel way. More importantly, contrary to decomposition techniques based on the first order necessary conditions, this let us see how we could generalize this approach to problems involving non-linear models or non-quadratic cost which will be a subject for future research.

Appendix A. Finite difference discretization of the augemented largangian algorithm.

A.1. Analysis of a 1D example. This section will mainly focus on the error induced by a finite difference discretization of the problem (4.1) on a simple 1D example. Assuming $\Omega = (-1, 1)$ and $\Gamma_\cap = \{0\}$, the optimizaton problem (4.1) now reads as

$$\min \sum_{i=1}^{2} \frac{1}{2}\|y^i\|^2_{L^2(\Omega_i))} + \frac{\alpha}{2}\|f^i\|^2_{L^2(\Omega_i)} + \lambda(y^1(0) - y^2(0)) + \frac{\rho}{2}(y^1(0) - y^2(0))^2$$
$$\text{t.q.} \begin{cases} -\dfrac{d^2}{dx^2}y^i = f^i, i = 1, 2, \\ \dfrac{d}{dx}y^i(0) = g, i = 1, 2, \\ u^1(-1) = 0, \ u^2(1) = 0. \end{cases}$$

We solve this problem with a direct method, which means that we will discretize directly the optimization problem. We use a second order 3-points centered finite difference scheme which yields

$$
\min \sum_{i=1}^{2} \frac{1}{2}\|y^i\|_{L^2} + \frac{\alpha}{2}\|f^i\|_{L^2}^2 + \frac{\lambda}{h}(y_N^1 - y_0^2) + \frac{\rho}{2h}(u_N^1 - y_0^2)^2
$$

$$
\text{(A.1)} \qquad \text{t.q.} \begin{cases} -\dfrac{2y_{N-1}^1 - 2y_N^1 + 2hg}{h^2} = f_N^1, \\ -\dfrac{2y_1^2 - 2y_0^2 - 2hg}{h^2} = f_0^2, \\ -\dfrac{y_{n-1}^1 - 2y_n^1 + y_{n+1}^1}{h^2} = f_n^1, \; n \in \{1, ..., N-1\} \\ -\dfrac{y_{n-1}^2 - 2y_n^2 + y_{n+1}^2}{h^2} = f_n^2, \; n \in \{1, ..., N-1\} \\ y_0^1 = 0, \; y_N^2 = 0. \end{cases}
$$

The lagrangian of (A.1) reads

$$
\begin{aligned}
\mathcal{L} =& \sum_{i=1}^{2} \frac{1}{2} \sum_{n=0}^{N} (y_n^i)^2 + \frac{\alpha}{2} \sum_{n=0}^{N} (f_n^i)^2 + \frac{\lambda}{h}(y_N^1 - y_0^2) + \frac{\rho}{2h}(y_N^1 - y_0^2)^2 \\
&+ \frac{1}{2}\left(\frac{2y_{N-1}^1 - 2y_N^2 + 2hg}{h^2} + f_N^1\right) p_N^1 + \frac{1}{2}\left(\frac{2y_1^2 - 2y_0^2 - 2hg}{h^2} + f_0^2\right) p_0^2 \\
&+ \sum_{n=1}^{N-1} \left(\frac{y_{n-1}^1 - 2y_n^1 + y_{n+1}^1}{h^2} + f_n^1\right) p_n^1 + \sum_{n=1}^{N-1} \left(\frac{y_{n-1}^2 - 2y_n^2 + y_{n+1}^2}{h^2} + f_n^2\right) p_n^2
\end{aligned}
$$

Differentating $\mathcal{L}$ w.r.t. y_n^2 for $n > 1$ gives $\partial_{y_n^2}\mathcal{L} = y_i^2 + \frac{p_{n+1}^2 - 2p_n^2 + p_{n-1}^2}{h^2} = 0$. Now differentating $\mathcal{L}$ w.r.t. y_1^2 and y_0^2, we get:

$$
\partial_{y_1^2}\mathcal{L} = y_1^2 + \frac{p_0^2}{h^2} + \frac{p_2^2 - 2p_1^2}{h^2} = y_1^2 + \frac{p_2^2 - 2p_1^2 + p_0^2}{h^2} = 0,
$$

$$
\partial_{y_0^2}\mathcal{L} = y_0^2 + \frac{\rho}{h}(y_N^1 - y_0^2) + \frac{\lambda}{h} - \frac{p_0^2}{h^2} + \frac{p_1^2}{h^2} = y_0^2 + \frac{\rho}{h}(y_N^1 - y_0^2) + \frac{\lambda}{h} + \frac{p_1^2 - p_0^2}{h^2} = 0.
$$

This is summarised as the following discrete adjoint equations

$$
\text{(A.2)} \qquad \begin{aligned} &-\frac{p_1^2 - p_0^2}{h} = hy_0^2 + \rho(y_N^1 - y_0^2) + \lambda, \\ &-\frac{p_{n-1}^2 - 2p_n^2 + p_{n+1}^2}{h^2} = y_n^2, \; n \in \{1, ..., N-1\}. \end{aligned}
$$

With the same calculations, one finds

$$
\text{(A.3)} \qquad \begin{aligned} &-\frac{p_{n-1}^1 - 2p_n^1 + p_{n+1}^1}{h^2} = y_n^1, \; n \in \{1, ..., N-1\}, \\ &-\frac{p_N^1 - p_{N-1}^1}{h} = hy_N^1 - \rho(y_N^1 - y_0^2) - \lambda. \end{aligned}
$$

Eventually, we derive $\mathcal{L}$ w.r.t. f_n^i and g:

$$
\partial_{f_n^1}\mathcal{L} = \alpha f_n^1 + p_n^1 = 0, \; n < N, \quad \partial_{f_N^1}\mathcal{L} = \alpha f_N^1 + \frac{p_N^1}{2} = 0,
$$

$$
\partial_{f_n^2}\mathcal{L} = \alpha f_n^2 + p_n^2 = 0, \; n > 0 \qquad \partial_{f_0^2}\mathcal{L} = \alpha f_0^2 + \frac{p_0^2}{2} = 0,
$$

$$\partial_g \mathcal{L} = -h^{-1}p_N^1 + h^{-1}p_0^2 = 0 \implies p_N^1 = p_0^2$$

The main difference with a discretization of (4.3) appears in the discretization of the Neumann boundary condition in (A.2)-(A.3). Instead of having the usual second order scheme discretizing the Neumann condition at the virtual interface, we obtain a first order approximation which introduces a perturbation hy. As we can see in Appendix A.2, this deteriorates the convergence on λ (and therefore, the convergence of the whole algorithm), since λ is directly linked to the normal derivative of the adjoint p at the virtual interface (see subsection 3.1). One could argue that we could use instead an indirect method, which would consist instead in discretizing the adjoint equation (4.3). That would indeed give us a better approximation of λ, but the calculation above proves that it would not be the adjoint of the discretized optimization problem, leading to an inexact optimality condition of the discretized problem.

A.2. Numerical 2D example. We now run the same numerical tests as in section 5 but using this time a 5-points finite difference scheme of order 2. For small α, the algorithm using the finite difference discretization fails at recovering the target.

α	$\|y - y_{\text{target}}\|_\infty$	$\|f\|_\infty$	$\|g - g_{\text{target}}\|_\infty$	Cost	# iterations
10	$3.45 \cdot 10^{-4}$	$1.27 \cdot 10^{-3}$	$2.16 \cdot 10^{-3}$	$4.04 \cdot 10^{-6}$	17
1	$4.91 \cdot 10^{-4}$	$1.27 \cdot 10^{-2}$	$3.06 \cdot 10^{-3}$	$4.02 \cdot 10^{-5}$	17
$1 \cdot 10^{-2}$	$1.62 \cdot 10^{-2}$	1.25	0.1	$3.95 \cdot 10^{-3}$	17
$1 \cdot 10^{-4}$	0.62	48.63	3.86	0.15	20
$1 \cdot 10^{-6}$	0.99	78.42	6.23	0.25	23

Table A.1: Error on the solution using different values of α and a finite difference scheme. $k = 2$, $C = 1$, $h = 1.10^{-2}$.

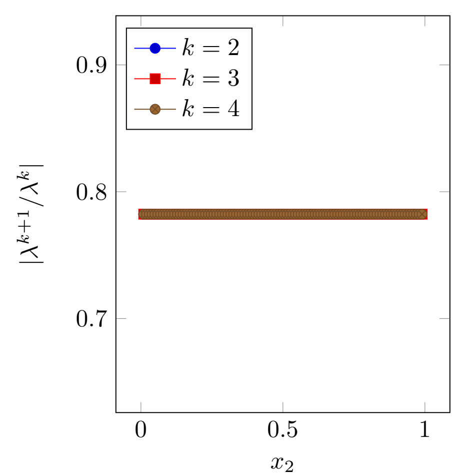


Fig. A.1: $|\lambda^{k+1}/\lambda^k|$ for different iterations, using finite differences.

Concerning now the rate of convergence on λ, the ratio between two iterations is much higher when compared to the finite element case, as shown in Figure A.1. This could be explained regarding the calculations given in Appendix A.1, and it shows that one must choose carefully the discretization used for these optimal control problem that are decomposed on subdomains.

REFERENCES

[1] Joe Alexandersen and Casper Schousboe Andreasen. A review of topology optimisation for fluid-based problems. *Fluids*, 5(1):29, 2020.

[2] Hedy Attouch and Mohamed Soueycatt. Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. applications to games, PDE's and control. *Pacific Journal of Optimization*, 5(1):17–37, 2008.

[3] Philip Avery and Charbel Farhat. The FETI family of domain decomposition methods for inequality-constrained quadratic programming: Application to contact problems with conforming and nonconforming interfaces. *Computer Methods in Applied Mechanics and Engineering*, 198(21-26):1673–1683, 2009.

[4] Maïtine Bergounioux. Augmented Lagrangian method for distributed optimal control problems with state constraints. *Journal of Optimization Theory and Applications*, 78(3):493–521, 1993.

[5] Maïtine Bergounioux and Mounir Haddou. A SQP-augmented Lagrangian method for optimal control of semilinear elliptic variational inequalities. In *Control and Estimation of Distributed Parameter Systems*, pages 57–72. Springer, 2003.

[6] Maïtine Bergounioux and Karl Kunisch. Augemented Lagrangian techniques for elliptic state constrained optimal control problems. *SIAM Journal on Control and Optimization*, 35(5):1524–1543, 1997.

[7] Dimitri P Bertsekas. Constrained optimization and Lagrange Multiplier methods. *Computer Science and Applied Mathematics*, 1982.

[8] Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[9] Aïcha Bounaim. *Méthodes de décomposition de domaine: Application à la résolution de problèmes de contrôle optimal*. PhD thesis, Université Joseph-Fourier-Grenoble I, 1999.

[10] Haïm Brézis. *Analyse fonctionnelle: théorie et applications*. Collection Mathématiques appliquées pour la maîtrise. Masson, 1983.

[11] Eduardo Casas. Control of an elliptic problem with pointwise state constraints. *SIAM Journal on Control and Optimization*, 24(6):1309–1318, 1986.

[12] Bastien Chaudet-Dumas and Jean Deteix. Shape derivatives for an augmented Lagrangian formulation of elastic contact problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:S14, 2021.

[13] Zhiming Chen and Jun Zou. An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. *SIAM Journal on Control and Optimization*, 37(3):892–910, 1999.

[14] Marco Discacciati, Paola Gervasio, and Alfio Quarteroni. The interface control domain decomposition (ICDD) method for elliptic problems. *SIAM Journal on Control and Optimization*, 51(5):3434–3458, 2013.

[15] Charbel Farhat and Francois-Xavier Roux. A method of finite element tearing and interconnecting and its parallel solution algorithm. *International journal for numerical methods in engineering*, 32(6):1205–1227, 1991.

[16] Martin J Gander. Optimized Schwarz methods. *SIAM Journal on Numerical Analysis*, 44(2):699–731, 2006.

[17] Martin J Gander and Hui Zhang. Schwarz methods by domain truncation. *Acta Numerica*, 31:1–134, 2022.

[18] Fernando Gaspoz, Christian Kreuzer, Andreas Veeser, and Winnifried Wollner. Quasi-best approximation in optimization with pde constraints. *Inverse Problems*, 36(1):014004, 2019.

[19] Wei Gong, Felix Kwok, and Zhiyu Tan. Convergence analysis of the Schwarz alternating method for unconstrained elliptic optimal control problems, 2022.

[20] Max Gunzburger and Jeehyun Lee. A domain decomposition method for optimization problems for partial differential equations. *Computers & Mathematics with Applications*, 40(2-3):177–192, 2000.

[21] Michael Hintermüller and Karl Kunisch. Feasible and noninterior path-following in constrained minimization with low multiplier regularity. *SIAM Journal on Control and Optimization*, 45(4):1198–1221, 2006.

[22] Michael Hinze, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, 2008.

[23] L.S. Hou and Jangwoon Lee. A Robin-Robin non-overlapping domain decomposition method for an elliptic boundary control problem. *International Journal of Numerical Analysis & Modeling*, 8(3), 2011.

[24] Kazufumi Ito and Karl Kunisch. The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces. *Mathematical programming*, 46(1):341–360, 1990.

[25] Kazufumi Ito and Karl Kunisch. The augmented Lagrangian method for parameter estimation

in elliptic systems. *SIAM Journal on Control and Optimization*, 28(1):113–136, 1990.

[26] Kazufumi Ito and Karl Kunisch. Augmented Lagrangian–SQP methods for nonlinear optimal control problems of tracking type. *SIAM journal on control and optimization*, 34(3):874–891, 1996.

[27] Alfredo Iusem and Rolando Gárciga Otero. Inexact version of proximal point and augmented lagrangian algorithms in Banach spaces. *Numerical Functional Analysis and Optimization*, 22(5-6):609–640, 2001.

[28] Christian Kanzow, Daniel Steck, and Daniel Wachsmuth. An augmented Lagrangian method for optimization problems in Banach spaces. *SIAM Journal on Control and Optimization*, 56(1):272–291, 2018.

[29] Phan Quoc Khanh, Tran Hue Nuong, and Michel Théra. On duality in nonconvex vector optimization in Banach spaces using augmented Lagrangians. *Positivity*, 3(1):49–64, 1999.

[30] Karl Kunisch and Xue-Cheng Tai. Nonoverlapping domain decomposition methods for inverse problems. In *Proceedings of 9th International Conference on Domain Decompostion Methods, Editors P. Bjorstard, M. Espedal and D. Keyes, John Wiley and Sons*. Citeseer, 1997.

[31] Jacques Louis Lions. *Optimal Control of Systems Governed by Partial Differential Equations:*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer-Verlag, 1971.

[32] Jacques Louis Lions and Enrico Magenes. *Non-homogeneous boundary value problems and applications: Vol. 1*, volume 181. Springer Science & Business Media, 2012.

[33] Frédéric Magoulès and François-Xavier Roux. Lagrangian formulation of domain decomposition methods: A unified theory. *Applied Mathematical Modelling*, 30(7):593–615, 2006. Parallel and Vector Processing in Science and Engineering.

[34] Jan Hendrik Maruhn. An augmented Lagrangian algorithm for optimization with equality constraints in Hilbert spaces. Master's thesis, Virginia Tech, 2001.

[35] Alfio Quarteroni and Alberto Valli. *Domain decomposition methods for partial differential equations*. Oxford University Press, 1999.

[36] Alexandre Vieira, Alain Bastide, and Pierre-Henri Cocquet. Topology optimization for steady-state anisothermal flow targeting solids with piecewise constant thermal diffusivity. *Applied Mathematics & Optimization*, 85(3):1–32, 2022.

[37] Alexandre Vieira and Pierre-Henri Cocquet. The Boussinesq system with non-smooth boundary conditions: existence, relaxation and topology optimization. *accepted for publication in SIAM Journal on Control and Optimization*, 2022.

[38] Eberhard Zeidler. *Nonlinear functional analysis and its applications: II/B: nonlinear monotone operators*. Springer Science & Business Media, 2013.