LiQuOFETI : a FETI-inspired method for elliptic quadratic optimal control problems

Alexandre Vieira [1] Pierre-Henri Cocquet [2]

[1]Univ. of Reunion Island [2]Univ. of Pau

Problem statement

The goal is to solve, with some decomposed scheme, the linear quadratic optimal control

$$\min \frac{1}{2}\int_\Omega (y - y_{\text{target}})^2 + \frac{\alpha}{2}\int_\Omega f^2$$
$$\text{s.t.} \begin{cases} \mathcal{A}y = -\text{div}\,(A(x)\nabla y) + \text{div}\,(b(x)y) + (c(x)+\mu)y = F + f, \\ y|_{\partial\Omega} = 0. \end{cases} \tag{1}$$

Here, we will assume that $A \in L^\infty(\Omega, \mathbb{R}^{d\times d})$, $b \in W^{1,\infty}(\Omega;\mathbb{R}^d)$, $c \in L^\infty(\Omega)$, $F \in L^2(\Omega)$ and $\mu > 0$.

A direct approach

In order to find the solutions of this problem, it is usual to solve necessary and sufficient conditions of optimality, expressed through the adjoint equation

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

The resolution of the coupled direct-adjoint equations through a decomposition technique has been analyzed (see for instance Gong et al. 2022). However, we find this approach restrictive for several reasons.

1. As with the indirect numerical methods for optimal control problems, focusing on resolving the necessary conditions of optimality turns out to be limited for numerous non-linear/non-quadratic optimization problems. This is mainly due to the fact that the resulting system of optimality is expressed as a DAE, which can be hardly solved, even without any decomposition. It may be even harder if you add further constraints on the state and/or the control, resulting in searching for the solution of variational inequalities with algebraic constraints.
2. Even if we focus on computing the gradient of the cost using only the direct and adjoint equations (and therefore, forgetting about the algebraic equation for a moment), it is still unclear how precise the computation of the states should be in order to compute an approximate gradient, that will then be used in a descent algorithm. The parallelization of such approach is also a source of questions.

Instead we will try a *direct* approach. As long as possible, we will stay in an optimization framework, and decompose directly in the constraints.

Theorem 1: Equivalent decomposed formulation

Problem (1) is equivalent to

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

where $\Gamma_\cap = \overline{\partial\Omega_1} \cap \overline{\partial\Omega_2}$.

Introduction of a virtual control

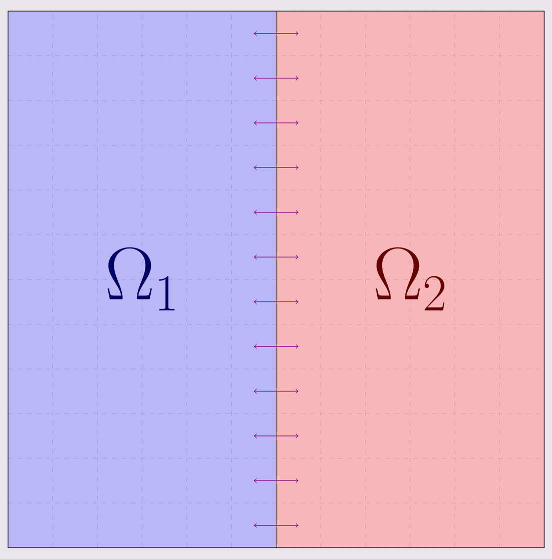


Figure 1. Decomposition idea.

The equivalence (2) boils down to two main ideas.

- The solution should be continuous at the interface $\Gamma_\cap$.
- The normal derivative $\partial_{n_\mathcal{A}} y_i$ becomes a new unknown that must be controlled with the same function, assuring the continuity of the normal derivative.

An augmented Lagrangian approach

The biggest challenge consists in finding a way to solve (2) with the continuity constraint $y_1|_{\Gamma_\cap} = y_2|_{\Gamma_\cap}$.

For this, we choose an augmented Lagrangian approach, and check its convergence. The new problem to solve now reads:

$$\min \frac{1}{2}\sum_{i=1}^{2} \|y_i - y_{\text{target}}\|^2_{L^2(\Omega_i)} + \alpha\|f_i\|^2_{L^2(\Omega_i)} + \int_{\Gamma_\cap} \lambda(y_1 - y_2) + \frac{\rho}{2}\int_{\Gamma_\cap}(y_1 - y_2)^2$$
$$\text{s.t.} \begin{cases} \mathcal{A}y_i = F + f_i & \text{in } \Omega_i, \\ y_i|_{\partial\Omega} = 0, \\ \partial_{\mathbf{n}_\mathcal{A}} y_i|_{\Gamma_\cap} = (-1)^{i+1} g, \; i = 1,2, \end{cases} \tag{3}$$

Note that (3) can be solved in a highly parallel framework, since the computation of the state and the update of the control can be done independently on each subdomain. Only the update of the virtual control g would need a synchronization.

References

Gong, W., F. Kwok, and Z. Tan (2022). *Convergence analysis of the Schwarz alternating method for unconstrained elliptic optimal control problems.* arXiv: 2201.00974 [math.NA].

Algorithm: update of the multiplier

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 (3) 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

Theorem 2: Convergence of the algorithm

Denote $x^k = (f_1^k, f_2^k, g^k)$ *the solutions produced by Algorithm 1, and suppose it converges to some* x^*. *Define* f^k *as* $f|_{\Omega_i} = f_i^k$, *and* y^k *the associated state. Then* f^k, y^k *converge to the solution of* (1).

A Fourier analysis of λ^k

Using a Fourier Analysis of the necessary and sufficient conditions of optimality of (3) when $F = 0$, $y_{\text{target}} = 0$ (analysis of the error correction), one may prove that the update of λ in the algorithm produces iterates $\{\lambda^k\}_k$ such that

$$R := \frac{\hat{\lambda}^{k+1}}{\hat{\lambda}^k} = \left(1 - i\rho^k\alpha^{-1/2}(D_+ - D_-)^{-1}\right)^{-1}$$

where $D_\pm = \sqrt{\pm i\alpha^{-1/2} + \omega^2}$, and $\hat{\lambda}$ is the Fourier transform of λ.
This coefficient shows how fast $\{\lambda^k\}$ converges, and it seems to converge fast!

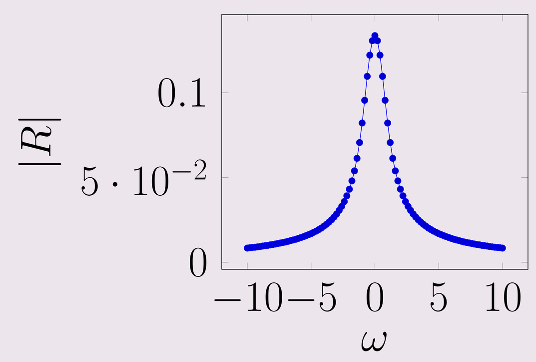
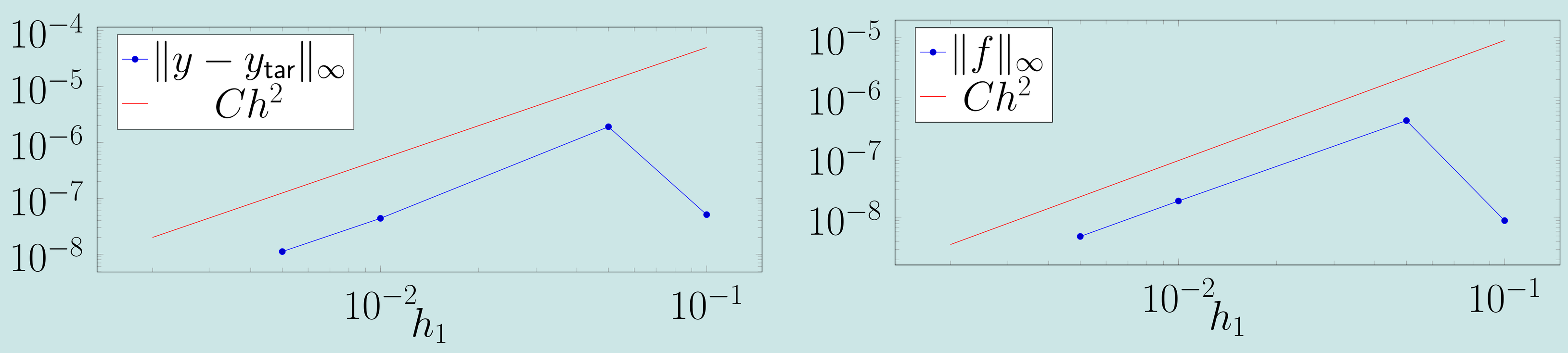


Figure 2. $|R|$ for $\rho^k = 3$, $\alpha = 1$.

Numerical example

We solve (1) with where $\Omega = [-1,1] \times [0,1]$, $y_{\text{target}}(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi x_2)$ and $F(x_1,x_2) = 8\pi^2\sin(2\pi x_1)\sin(2\pi x_2)$, $\alpha = 1$. The optimal solution is $f^* = 0$, $y^* = y_{\text{target}}$. We solve this problem using our augmented lagrangian method. The problem is discretized using Q1 elements on a structured uniform mesh, and the interface is placed at $\Gamma_\cap = \{0\} \times [0,1]$. We retrieve a second order convergence of the solution with respect to the discretization stepsize.



(a) Error on the state function. (b) Error on the control function.

We see also the fast covergence of λ in the case of $y_{\text{target}} = F = 0$, which is even better than predicted by the theory.

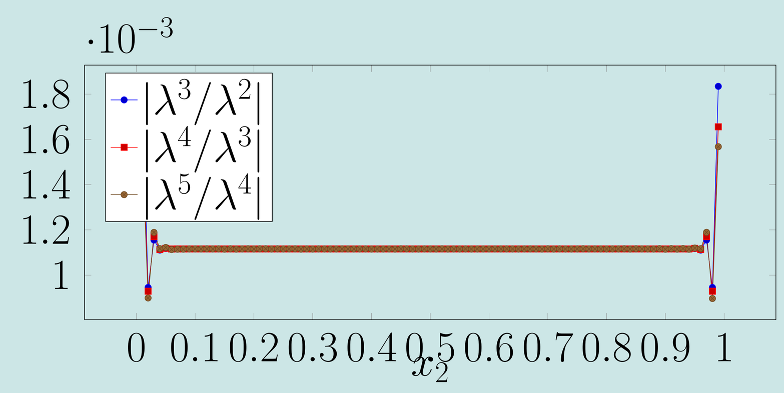


Figure 4. $|\lambda^{k+1}/\lambda^k|$ for different iterations.

alexandre.vieira@univ-reunion.fr pierre-henri.cocquet@univ-pau.fr