Nonoverlapping Domain Decomposition of Parabolic Optimal Control Problems Revisited: Decompose-then-Optimize versus Optimize-then-Decompose

Martin J. Gander[0000−0001−8450−9223] and Günter Leugering[0000−0002−1086−991X]

1 Introduction

Non-overlapping domain decomposition methods for PDE-constrained optimization problems have a long history. The early treatment by Benamou and Després [1] focused on the extension of the P.L. Lions algorithm for elliptic problems [15] to Helmholtz problems and then elliptic optimal control problems using a complex formulation. For parabolic optimal control problems, Lee [12] used an artificial control (that we would now call a virtual control) at the interface and a penalization of the mismatch of the states at the interface. Heinkenschloss and Herty [10] used a semi-discrete setting to apply a Neumann-type DD-technique. For more recent contributions, see e.g. Leugering [13] and Leugering et al. [14] for fractional diffusion equations on networks. We do not dwell here on the rich literature on time-domain decomposition of parabolic optimal control problems, for examples, see Gander et al. [6, 5] , and for hyperbolic problems, we refer to the monograph [11] and also [3].

2 Augmented Lagrange approach for virtual controls at the interface

We consider a standard optimal control problem for a parabolic equation in $\Omega\times(0,T)$, where $\bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2$ with common interface $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$. We introduce the pivot space $H = L^2(\Omega)$, the energy space $V = H_0^1(\Omega)$, as well as the domain space

Martin Gander
Université de Genève Section de Mathématiques,Rue du Conseil-Général 9, CP 64 1211 Genève 4, Suisse e-mail: martin.gander@unige.ch

Günter Leugering
Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Cuerstraß e 11, 91058 Erlangen. e-mail: guenter.leugering@fau.de

$D(A) := H^2(\Omega) \cap H_0^1(\Omega)$ and refer to the Gelfand triple $V \subset H \subset V^*$, the setting in which the mathematical analysis takes place. The set of distributed controls $U = H$ represents unconstrained distributed controls , and the problem reads

$$\begin{aligned}
&\min_{u,y} J^{\kappa,\nu}(u,y) := \frac{\kappa}{2}\int_0^T\int_\Omega \|y-y^d\|^2 dxdt + \frac{\nu}{2}\int_0^T\int_\Omega \|u\|^2 dxdt \quad \text{s.t.} \qquad (1)\\
&\partial_t y - \Delta y = u, \quad \text{in } Q := \Omega\times(0,T),\\
&y = 0 \ \text{ on } \partial\Omega\times(0,T),\\
&y(\cdot,0) = y^0 \ \text{ in } \Omega.
\end{aligned}$$

The problems of well-posedness in terms of so-called weak-, mild- and strong solutions and the characterization of the corresponding optimality system are standard an can be looked up in the text-book literature. We, therefore, just write down the optimality system corresponding to (1),

$$\begin{aligned}
&\partial_t y - \Delta y = \frac{1}{\nu}p, \quad \text{in } Q := \Omega\times(0,T),\\
&\partial_t p + \Delta p = \kappa(y-y^d), \quad \text{in } Q := \Omega\times(0,T),\\
&y = 0,\ \ p = 0 \ \text{ on } \partial\Omega\times(0,T),\\
&y(\cdot,0) = y^0, p(\cdot,T) = 0 \ \text{ in } \Omega.
\end{aligned} \tag{2}$$

The problem that we discuss in these notes concerns the decomposition of the optimal control problem (1) or the optimality system (2) with respect to the domains Ω_1, Ω_2 in terms of iterative non-overlapping domain decomposition procedures. Dealing with the optimality system as the object of decomposition can be framed as *optimize-then-decompose*, since the optimization problem has been replaced be the optimality system before considering decomposition with respect to the domains. The question then is whether it is possible to design a domain decomposition of the original optimal control problem (1) which would then be an example of the *decompose-then-optimize* principle. This question has been posed by the first author during his lecture at DD28, and he also provided a novel algorithm in that respect for an elliptic optimal control problem, see [4]. We would like to put this approach for parabolic problems into the context of *virtual controls*, a context that has been established in the work of J.L. Lions and O. Pironneau [16] and that has been used intensively thereafter. An equivalent formulation of the equality constraints in (1) is given by

$$\begin{aligned}
&\partial_t y_i - \Delta y_i = u_i, \quad \text{in } Q_i := \Omega_i\times(0,T),\\
&\partial_{n_1} y_1 + \partial_{n_2} y_2 = 0,\ \ y_1 = y_2, \quad \text{on } \Gamma\times(0,T),\\
&y_i = 0 \ \text{ on } \partial\Omega\setminus\Gamma\times(0,T),\\
&y_i(\cdot,0) = y_i^0 \ \text{ in } \Omega_i,\ \ i = 1,2.
\end{aligned} \tag{3}$$

The idea is to introduce virtual controls g_i satisfying $g_1 + g_2 = 0$ such that the transmission conditions $\partial_{n_1} y_1 + \partial_{n_2} y_2 = 0$ are satisfied and the controls are used to eliminate the discrepancy between y_1, y_2 on the interface, i.e. to fulfill the state constraints $y_1 - y_2$ on $\Gamma \times (0,T)$. The algorithm in [4] for elliptic problems consists in relaxing the continuity condition, i.e. the linear state constraint, by an augmented Langrangian approach as follows: we first extend the cost function by an augmented Lagrangian relaxation of the constraints,

$$J^{\kappa,\nu,\rho}(u,g,y;\lambda) := J^{\kappa,\nu}(u,y) + \int_{\Gamma\times(0,T)} \lambda(y_1 - y_2) + \frac{\rho}{2}|y_1 - y_2|^2 d\gamma dt. \quad (4)$$

The control set $U := \{(u,g) \in L^2(0,T;\Omega\times\Gamma)\}|g_1 + g_2 = 0)\}$ reflects the control constraint for the virtual controls g_i. The admissible set then is $\Sigma = \{(u,g,y)|(u,g) \in U, y \text{ s.t. } y \text{ satisfies (5)}\}$, which is

$$\begin{aligned} &\partial_t y_i - \Delta y_i = u_i, \quad \text{in } Q_i := \Omega_i \times (0,T), \\ &\partial_{n_i} y_i = g_i, \quad \text{on } \Gamma \times (0,T), \\ &y_i = 0 \quad \text{on } \partial\Omega \times (0,T), \\ &y_i(\cdot,0) = y_i^0 \quad \text{in } \Omega_i,\ i = 1,2. \end{aligned} \quad (5)$$

The constrained saddle-point problem then reads

$$\min_{(u,g,y)\in\Sigma} \max_{\lambda} J^{\kappa,\nu,\rho}(u,g,y;\lambda).$$

In fact, the problem is then treated in the so-called reduced formulation set-up, where the solutions $y_i(u_i,g_i) = S_i(u_i,g_i)$ are given by the solution operator,

$$\min_{(u,g)\in U} \max_{\lambda} J^{\kappa,\nu,\rho}(u,g,S(u,g);\lambda). \quad (6)$$

The Hestenes-type saddle-point iteration for (6) consists in gradient-descent steps for $J^{\kappa,\nu,\rho}$ followed by updating rules for the Lagrange multiplier λ and the penalty parameter ρ. Due to space limitations, we omit the details. Rather, we now have a closer look at the role of the constrained virtual controls. We first remark that, according to the classical work of Glowinski and LeTallec [9], we may introduce an extra variable q in $H_\Gamma := L^2(\Gamma \times (0,T))$ such that $y_i = q, i = 1,2$. Moreover, we may also introduce another variable $z \in H_\Gamma$ such that $g_1 = z = -g_2$ on $\Gamma \times (0,T)$, hence the control constraint $g_1 + g_2 = 0$ is satisfied. We then introduce an augmented Lagrange relaxation of the two constraints. To this end, we introduce the function

$$J^{\kappa,\nu,\rho,\sigma}(u,g,y,q,z;\lambda,\eta) := \sum_{i=1}^{2}\int_{\Omega_i}\left(\frac{\kappa}{2}|y_i - y_i^d|^2 + \frac{\nu}{2}|u_i|^2\right)dxdt \tag{7}$$
$$+\sum_{i=1}^{2}\int_0^T\int_\Gamma\left(\eta_i((-1)^i g_i - z) + \frac{\sigma}{2}|(-1)^i g_i - z|^2\right)d\gamma dt$$
$$+\sum_{i=1}^{2}\int_0^T\int_\Gamma\left(\lambda_i(y_i - q) + \frac{\rho}{2}|y_i - q|^2\right)d\gamma dt.$$

We then look for a saddle-point of the problem

$$\min_{(u,g,y)\in\Sigma,(q,z)\in H^2_\Gamma}\ \max_{\lambda,\eta} J^{\kappa,\nu,\rho,\sigma}(u,g,y,q,z;\lambda,\eta). \tag{8}$$

As in primal-dual optimization, we introduce the Lagrangian

$$\mathcal{L}(u,g,q,z,y;\lambda,\eta,p) := J^{\kappa,\nu,\rho,\sigma}(u,g,y,q,z;\lambda,\eta) \tag{9}$$
$$+\sum_{i=1}^{2}\int_0^T\int_{\Omega_i}\left(\partial_t y_i p_i + \nabla y_i \nabla p_i - u_i p_i\right)dxdt.$$

In order to solve the corresponding saddle-point problem iteratively, we invoke the Uzawa-type algorithm given by Glowinski and LeTallec [9][Alg.3] which is a fractional step algorithm,where $\rho^k, \sigma^k > 0$, $k = 0, 1, \ldots$ are given numbers.

Algorithm 1. *1. Given $q^{k-1}, z^{k-1}, \lambda^k, \eta^k$,*
2. Solve for u^k, g^k, y^k and p^k the equation

$$\partial_{u,g,y,p}\mathcal{L}(u,g,y,,q^{k-1},z^{k-1}p,\lambda^k,\eta^k) = 0.$$

3. Update λ^k and η^k (fractional step)

$$\lambda_i^{k+\frac{1}{2}} = \lambda_i^k + \rho^k(y_i^k - q^{k-1}),\quad \eta_i^{k+\frac{1}{2}} = \eta_i^k + \sigma^k((-1)^i g_i^k - z^{k-1}).$$

4. Solve for q^k and z^k the equation

$$\partial_{q,z}\mathcal{L}(u^k,g^k,y^k,q,z,p^k,\lambda^{k+\frac{1}{2}},\eta^{k+\frac{1}{2}}) = 0.$$

5. Complete the update of λ^{k+1} and η^{k+1},

$$\lambda_i^{k+1} = \lambda_i^{k+\frac{1}{2}} + \rho^k(y_i^k - q^k),\quad \eta_i^{k+1} = \eta_i^{k+\frac{1}{2}} + \sigma^k((-1)^i g_i^k - z^k).$$

6. Check the convergence and stop if satisfied, otherwise increase $k \to k+1$ and return to 1.

Remark 2 It should be noticed that in steps 2. and 4. one can use a gradient procedure instead of solving the corresponding optimality systems. As in [4], we can resort to the reduced approach, where one uses the solution operator, denoted by $y(u, g)$,

$$J_{red}^{\kappa,\nu,\rho,\sigma}(u, g, q, z; \lambda, \eta) := \sum_{i=1}^{2} \int_{\Omega_i} \left(\frac{\kappa}{2}|y_i(u, g) - y_i^d|^2 + \frac{\nu}{2}|u_i^2|\right) dxdt \qquad (10)$$
$$+ \sum_{i=1}^{2} \int_0^T \int_\Gamma \left(\eta_i((-1)^i g_i - z) + \frac{\sigma}{2}|(-1)^i g_i - z|^2\right) d\gamma dt$$
$$+ \sum_{i=1}^{2} \int_0^T \int_\Gamma \left(\lambda_i(y_i(u, g) - q) + \frac{\rho}{2}|y_i(u, g) - q|^2\right) d\gamma dt.$$

In particular, as seen in Glowinski and LeTallec [9, Algorithm 3], one can replace step 2. by

Find u_i^k, g_i^k s.t. $\mathcal{J}_{red}^{\kappa,\rho,\sigma}(u^k, g^k, q^{k-1}, z^{k-1}, \lambda^k, \eta^k) \leq \mathcal{J}_{red}^{\kappa,\rho}(u, g, q^{k-1}, z^{k-1}, \lambda^k, \eta^k)$,

$\forall (u, g)$ admissible , while step 4. is replaced by

Find q^k, z^k s.t. $\mathcal{J}_{red}^{\kappa,\rho,\sigma}(u^k, g^k, q^{k-1}, z^{k-1}, \lambda^k, \eta^k) \leq \mathcal{J}_{red}^{\kappa,\rho}(u^k, g^k, q, z, \lambda^k, \eta^k)$,

$\forall (q, z)$ admissible . The (reduced) algorithm above can thus also be used in the *decompose-then-optimize* framework, and adds to the new procedure from [4].

We are however going to show now that for a special choice of ρ^k, σ^k, namely $\rho^k = \rho, \sigma^k = \sigma = \frac{1}{\rho}$, Algorithm 1 can be interpreted as an *optimize-then-decompose* method, i.e. a domain decomposition method for the global optimality system (2). To do so, we go back to Algorithm 1, with the special choice of ρ^k, σ^k, and derive for (9) the optimality conditions, namely

$$\begin{aligned}
&\partial_t y_i - \Delta y_= u_i, \ \text{ in } Q := \Omega_i \times (0, T),\\
&\partial_t p_i + \Delta p_i = \kappa(y_i - y_i^d), \ \text{ in } Q := \Omega_i \times (0, T),\\
&y_i = 0, \ p_i = 0 \ \text{ on } \partial\Omega \setminus \Gamma \times (0, T),\\
&\partial_{n_i} y_i = g_i, \ \partial_{n_i} p_i = -\lambda_i - \rho(y_i - q), \ \text{ on } \Gamma \times (0, T),\\
&y(\cdot, 0) = y^0, p(\cdot, T) = 0 \ \text{ in } \Omega,
\end{aligned} \qquad (11)$$

for the state and the adjoint state, and

$$\begin{aligned} u_i &= \frac{1}{\nu} p_i, \quad \text{in } \Omega_i \times (0,T), \\ g_i &= \rho p_i - (-1)^i (\rho \eta_i - z) \quad \text{on } \Gamma \times (0,T), \\ q &= \frac{1}{2\rho} \sum_{i=1}^{2} \lambda_i + \frac{1}{2} \sum_{i=1}^{2} y_i \quad \text{on } \Gamma \times (0,T), \\ z &= \frac{1}{2} \sum_{i=1}^{2} (-1)^i g_i + \frac{\rho}{2} \sum_{i=1}^{2} \eta_i \quad \text{on } \Gamma \times (0,T), \end{aligned} \tag{12}$$

for the controls (virtual and real). We now solve, in step 2. of Algorithm 1, the system (11) for (y_i^k, p_i^k) given the values of $q^{k-1}, z^{k-1}, \lambda^k, \eta^k$. After the fractional step 3. of Algorithm 1, we obtain

$$q^k = \frac{1}{2\rho} \sum_{i=1}^{2} \lambda_i^{k+\frac{1}{2}} + \frac{1}{2} \sum_{i=1}^{2} y_i^k, \quad z^k = \frac{1}{2} \sum_{i=1}^{2} (-1)^i g_i^k + \frac{\rho}{2} \sum_{i=1}^{2} \eta_i^{k+\frac{1}{2}}. \tag{13}$$

The major step now is the update 5., where we use step 3. and (13),

$$\lambda_i^{k+1} = \lambda_i^{k+\frac{1}{2}} + \rho(y_i^k - q^k) = \left(\lambda_i^k - \frac{1}{2} \sum_{i=1}^{2} \lambda_j^k \right) + 2\rho \left(y_i^k - \frac{1}{2} \sum_{i=1}^{2} y_j^k) \right).$$

This shows that $\lambda_i^{k+1} = -\lambda_j^{k+1}$. Similarly, we obtain

$$\eta_i^{k+1} = \eta_i^{k+\frac{1}{2}} + \frac{1}{\rho}((-1)^i g_i^k - z^k) = \left(\eta_i^k - \frac{1}{2} \sum_{i=1}^{2} \eta_j^k \right) + \frac{2}{\rho} \left((-1)^i g_i^k - \frac{1}{2} \sum_{i=1}^{2} (-1)^j g_j^k) \right),$$

which shows similarly that also $\eta_i^{k+1} = -\eta_j^{k+1}$.

We can now move to the next step $k+1$. We obtain y_i^{k+1}, p_i^{k+1} and, using the last identity

$$\begin{aligned} g_i^{k+1} &= \rho p_i^{k+1} - (-1)^i \left(-\rho \eta_j^{k+1} - z^k \right) \\ &= \rho p_i^{k+1} - (-1)^i \left(-\rho \eta_j^k - 2(-1)^j g_j^k + z^{k-1} \right) \\ &= \rho p_i^{k+1} - (-1)^j \left(\rho \eta_j^k - z^{k-1} \right) - 2 g_j^k \\ &= \rho p_i^{k+1} - \rho p_j^k - g_j^k. \end{aligned}$$

We therefore obtain for the normal derivatives

$$\begin{aligned} &\partial_{n_i} y_i^{k+1} = g_i^{k+1} = \rho p_i^{k+1} - \rho p_j^k - \partial_{n_j} y_j^k \quad \Longleftrightarrow \\ &\partial_{n_i} y_i^{k+1} - \rho p_i^{k+1} = -\rho p_j^k - \partial_{n_j} y_j^k. \end{aligned} \tag{14}$$

We now look at the interface condition for p_i^{k+1},

$$
\begin{aligned}
\partial_{n_i} p_i^{k+1} &= -\lambda_i^{k+1} - \rho(y_i^{k+1} - q^k) \\
&= -\rho y_i^{k+1} + \rho y_j^k + \lambda_j^k + \rho(y_j^k - q^{k-1}) \\
&= \rho y_i^{k+1} + \rho y_j^k - \partial_{n_j} p_j^k,
\end{aligned}
\tag{15}
$$

which shows that

$$
\partial_{n_i} p_i^{k+1} + \rho y_i^{k+1} = \rho y_j^k - \partial_{n_j} p_j^k.
$$

We can now rewrite the local optimality system (11) at step $k+1$ using (14), (15) as

$$
\begin{aligned}
&\partial_t y_i^{k+1} - \Delta y_i^{k+1} = \frac{1}{\nu} p_i^{k+1}, \;\; \text{in } Q := \Omega_i \times (0,T), \\
&\partial_t p_i^{k+1} + \Delta p_i^{k+1} = \kappa(y_i^{k+1} - y_i^d), \;\; \text{in } Q := \Omega_i \times (0,T), \\
&y_i^{k+1} = 0, \;\; p_i^{k+1} = 0 \;\; \text{on } \partial\Omega \setminus \Gamma \times (0,T), \\
&\partial_{n_i} y_i^{k+1} - \rho p_i^{k+1} = -\rho p_j^k - \partial_{n_j} y_j^k, \\
&\partial_{n_i} p_i^{k+1} + \rho y_i^{k+1} = \rho y_j^k - \partial_{n_j} p_j^k \;\; \text{on } \Gamma\times, (0,T) \\
&y(\cdot,0) = y^0, p(\cdot,T) = 0 \;\; \text{in } \Omega.
\end{aligned}
\tag{16}
$$

This leads to the following important remarks:

- The system (16) corresponds precisely to the non-overlapping domain decomposition method suggested by P.L. Lions for elliptic problems.
- The iteration (16) is, thus, also the result of a non-overlapping domain decomposition method applied directly to the global optimality system (2), which puts the method into the *optimize-then-decompose* framework.
- Based on the derivation of (16) from the augmented Lagrangian saddle-point iteration, however, we see that only the local optimality system (11) in the step 2. of Algorithm 1 comes into play, while the decomposition is governed by the updates of the Lagrange multipliers and the virtual controls at the interface. In this sense, the iteration can also be seen as a *decompose-then-optimize* technique, in particular if one relaxes the computation of the solution of (11) to gradient descent steps as indicated in Remark 2.
- The proof of convergence of the method can be done along the lines of e.g. [13], [14].
- The optimal choice of the parameters σ and ρ has been discussed already in [7][Thm. 5.4], where convergence rates are given, and [8].
- In [11], under-relaxation of a similar scheme in the context of PDE-constrained optimal control problems has been discussed.
- Complexitiy and performance of augmented Lagrangian algorithms have been considered for instance in [2].
- The method can be extended to the parabolic p-Laplace equation (see [13]). Moreover, currently for 1-d problems, we can extend the method to space-time-fractional parabolic problems on graphs, which, in fact, was the subject of the presentation at DD28 of the second author.

- We also note that given the iteration at step k, the system (16) can be interpreted as the optimality system of a local optimization problem on the domain Ω_i. Due to space limitations, we can not elaborate on this feature further and refer to [13].

References

1. Benamou, J.-D., Desp,rès, B.: A domain decomposition method for the Helmholtz equation and related optimal control problems, J. Comput. Phys., **136**(1), 68–82 (1997)
2. Birgin, E. G. and Martínez, J. M.: Complexity and performance of an Augmented Lagrangian algorithm, Optimization Methods and Software, 35(5), 885–920, (2020)
3. Bui, Duc-Quang, Delourme, B., Halpern, L., Kwok, F.:Optimized Schwarz Method in Time for Transport Control. In: International Conference on Domain Decomposition Methods, 93–100, Springer (2022)
4. Cocquet, P.-H., Gander, M. J.,Vieira, A.: Decompose-then-Optimize: a new approach to design domain decomposition methods for optimal control problems. Submitted (2024)
5. Gander, M. J.,Lu, Liu-Di: New time domain decomposition methods for parabolic control problems I: Dirichlet-Neumann and Neumann-Dirichlet algorithms. SINUM, in print, (2024)
6. Gander, M. J., Kwok, F.: Schwarz methods for the time-parallel solution of parabolic control problems. In: Domain decomposition methods in science and engineering XXII, Lecture Notes in Computational Science and Engineering, **104**, 207–216, Springer, Cham, (2016)
7. Fortin, M, Glowinski, R.: Augemtned Lagrangian Mehtods: Applications to the numerical solutions of boundary value problems, North-Holland Amsterdam-New York-Oxford, (1983)
8. Glowinski, R., Le Tallec, P.: Augmented Lagrangians and Operator Splitting in Nonlinear Mechanics, SIAM, Philadelphia,(1989)
9. Glowinski, R., Le Tallec, P.: Augmented Lagrangian interpretation of the nonoverlapping Schwarz alternating method. In: Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989), 224–231,SIAM, Philadelphia, PA, (1990)
10. Heinkenschloss, M., Herty, M.: A spatial domain decomposition method for parabolic optimal control problems. J. Comput. Appl. Math., **201**(1), 88-111 (2007)
11. Lagnese, J. E., Leugering, G.: Domain decomposition methods in optimal control of partial differential equations. International Series of Numerical Mathematics **148**, xiv+443 p. Birkhäuser Verlag, Basel (2004)
12. Lee, Jeehyun:An optimization-based domain decomposition method for parabolic equations. Appl. Math. Comput., **175**(2), 1644–1656 (2006)
13. Leugering, G.: Nonoverlapping domain decomposition and virtual controls for optimal control problems of p-type on metric graphs. In: Trélat, E., Zuazua, E. (eds), Handbook of Numerical Analysis,Chapter 7, Elesevier (2023), url= https://doi.org/10.1016/bs.hna.2022.11.002
14. Leugering, G., Mehandiratta, V., Mehra, M.: Non-Overlapping Domain Decomposition for 1D Optimal Control Problems Governed by Time-Fractional Diffusion Equations on Coupled Domains: Optimality System and Virtual Controls. Fractal and Fractional, **8**(3), (2024)
15. Lions, P.-L.: On the Schwarz alternating method. III. A variant for nonoverlapping subdomains. In: Third International Symposium on Domain Decomposition: Methods for Partial Differential Equations (Houston, TX, 1989), 202-223 SIAM, Philadelphia, PA, (1990)
16. Lions, J.-L., Pironneau, O.: Algorithmes parallèles pour la solution de problèmes aux limites. C. R. Acad. Sci. Paris Sér. I Math., **327**(11), 947–952 (1998)