Decompose-then-Optimize versus Optimize-then-Decompose for the Poisson Problem in Minimization Form

Pierre-Henri Cocquet[1], Martin J. Gander[2]

1 Introduction

The main motivation of this paper comes from our recent work [3] where we introduced the *decompose-then-optimize* (DtO) technique to solve PDE-constrained optimization problems, see also [8]. The idea is to rewrite the PDE constraint valid on the whole domain as an equivalent system of PDEs on subdomains only, coupled on the artificial interfaces by specific *continuity constraints*, which are then handled by various optimization techniques. Due to the many optimization techniques, one can discover systematically new DD methods doing this; the classical FETI method [4] is a prime example.

In contrast, the classical *optimize-then-decompose* (OtD) approach is to directly apply Domain Decomposition Methods (DDMs) to the first order optimality conditions given by the coupled direct and adjoint PDEs. This is well explored and leads to interesting new results for DDMs, see [6] for Schwarz methods, [7] for substructuring methods, [16] for non-overlapping DDMs, and [9] for Dirichlet-Neumann and Neumann-Neumann methods for elliptic control problems. In OtD, the resulting DDM iterations can also often be interpreted as solving optimal control problems on the subdomains, see e.g. [1, 2, 6, 7, 13, 15, 18].

A very classical question in the numerical analysis of optimal control problems is whether one should discretize-then-optimize, or better optimize-then-discretize [12]. There are advantages to both approaches, and in some cases, the two approaches also commute. We are precisely interested in the same question for DtO and OtD: we want to know in which cases DtO leads to a *known* DDM (of Schwarz, substructuring, Dirichlet-Neumann or Neumann-Neumann type) applied to the first-order optimality conditions of the PDE-

(1) Université de Pau et des Pays de l'Adour/ E2S-UPPA, SIAME, PAU, 64000, France. *P-H Cocquet is supported by the ANR Project DGMod number ANR-23-CE56-0007.* (2) Université de Genève, 24 rue de Général-Dufour, 1211 Genève 4, Switzerland, {martin.gander@unige.ch}{pierre-henri.cocquet@univ-pau.fr}

constrained optimization problem in the OtD approach, and when not. We provide specific examples of positive and negative answers to this question by considering the for simplicity Poisson equation

$$-\Delta y = f \quad \text{in } \Omega,\ y \in H_0^1(\Omega), \tag{1}$$

where Ω is a bounded open set of $\mathbb{R}^d$ with Lipschitz boundary. To fit the Poisson equation (1) into our DtO framework, we need to consider its equivalent form as a minimization problem

$$\min_{y\in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla y|^2 - f\, y\, dx. \tag{2}$$

For our investigation, it suffices to consider a non-overlapping decomposition of $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$ with common interface $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$. The DtO counterpart of (2) is then obtained by adding continuity constraints on Γ which gives

$$\begin{aligned} &\min_{y_i\in H^1(\Omega_i)} \sum_{i=1}^{2} \int_{\Omega_i} \frac{1}{2}|\nabla y_i|^2 - f\, y_i\, dx, \\ &y_i|_{\partial\Omega_i\setminus\Gamma} = 0,\ y_1|_\Gamma = y_2|_\Gamma. \end{aligned} \tag{3}$$

The solution of (3) can be computed using various optimization techniques, for example the augmented Lagrangian method, and each technique yields a potentially new DD method, FETI being one example.

In addition to the FETI method, the DtO approach has already been used by Glowinski in [10] and Le Tallec in [14], where the solution to problems of the form (2) has been computed using two Lagrange multipliers and a virtual control q for the constrains $y_1|_\Gamma = q = y_2|_\Gamma$, using an augmented Lagrangian algorithm with fractional steps, and it has been proved that the resulting iteration is exactly the non-overlapping DD method of Schwarz type with Robin transmission conditions introduced by P.L. Lions. Therefore, this method satisfies DtO=OtD.

In this paper, we show that other choices for the Lagrangian associated to (3) and different augmented Lagrangian algorithms yield methods for which DtO≠OtD, and we focus on the following three cases:

1. one multiplier for the constraint $y_1|_\Gamma = y_2|_\Gamma$ and the augmented Lagrangian method of Hestenes and Powell (see [11, 17]);
2. two multipliers and a virtual control q for the constraints $y_i|_\Gamma = q$ and the augmented Lagrangian method of Hestenes and Powell;
3. two multipliers and a virtual control q for the constraints $y_i|_\Gamma = q$ and the augmented Lagrangian method with fractional steps from [10, 14].

We explain why the two first cases lead to DtO≠OtD, and the third case satisfies DtO=OtD. We also perform a Fourier analysis for each method to get insight on the convergence behavior of the three methods and compare their performance.

2 Optimization-based DDM for which DtO≠OtD

We recall in Algorithm 1 the method of Hestenes [11] and Powell [17] to compute saddle-points of a general augmented Lagrangian $\mathcal{L}_\rho(y,\lambda)$.

Algorithm 1 : Multiplier method of Hestenes and Powell (1969)
1. Given λ_i^{k-1}, compute y_i^k such that
$$y_i^k = \arg\min_{y_i} \mathcal{L}_\rho(y_i^k;\lambda_i^{k-1}).$$
2. Update the multipliers as
$$\lambda_i^k = \lambda_i^{k-1} + \rho\partial_{\lambda_i}\mathcal{L}_\rho(y_i^k;\lambda_i^{k-1}).$$

We apply Algorithm 1 to compute the saddle point of two different augmented Lagrangians associated to (3):

1. One multiplier and no-virtual control: We consider

$$\mathcal{L}_\rho^{(1)}(y_i;\lambda) = \sum_{i=1}^{2}\int_\Omega \frac{1}{2}|\nabla y_i|^2 - f\,y_i\,dx + \int_\Gamma \lambda(y_1-y_2) + \frac{\rho}{2}|y_1-y_2|^2\,d\sigma$$

as augmented Lagrangian associated to (3). Using Algorithm 1 to compute the saddle point of $\mathcal{L}_\rho^{(1)}$, given λ^{k-1}, we first compute y_i^k minimizing $\mathcal{L}_\rho(y;\lambda_i^{k-1})$. Since the optimization problem is convex, we have to solve $\partial_{y_i}\mathcal{L}_\rho(y;\lambda_i^{k-1})=0$ i.e.

$$-\Delta y_i^k = f_i \text{ in } \Omega_i,\ y_i^k|_{\partial\Omega_i\setminus\Gamma}=0,\ \partial_{n_i}y_i^k = (-1)^i\left(\lambda^{k-1}+\rho(y_1^k-y_2^k)\right) \text{ on } \Gamma. \tag{4}$$

Since the right hand sides match in the second equation, the Neumann conditions match at each iteration of the algorithm, like in classical FETI methods. The multiplier is then updated by setting

$$\lambda^k = \lambda^{k-1} + \rho\left(y_1^k - y_2^k\right) \text{ on } \Gamma. \tag{5}$$

Using this relation in (4) we get $\partial_{n_1}y_1^k = -\lambda^k$ and $\partial_{n_2}y_2^k=\lambda^k$. This also holds at iteration $k-1$, and substituting $\lambda^{k-1} = -\partial_{n_1}y_1^{k-1}$ and $\lambda^{k-1}=\partial_{n_2}y_2^{k-1}$ into (4), we get on Γ, using that the normal derivatives have opposite signs, the two transmission conditions

$$\partial_{n_1}y_1^k + \rho y_1^k = \partial_{n_1}y_2^{k-1} + \rho y_2^k, \quad \partial_{n_2}y_2^k + \rho y_2^k = \partial_{n_2}y_1^{k-1} + \rho y_1^k. \tag{6}$$

If the y_i terms multiplied by ρ on the right were at iteration index $k-1$, this would be the classical non-overlapping Schwarz method of Lions with Robin transmission conditions. With iteration index k however, this method may not be interpreted as a *known* non-overlapping DDM, and hence DtO $\neq$ OtD in this case. In fact, the subproblems with transmission conditions (6) are not decoupled at iteration k and can thus not be solved independently, in contrast to DDMs!

To get more insight on the convergence properties of algorithm (4)-(5), we now perform a Fourier analysis on the error equation (i.e. $f = 0$) and in the special case where $\Omega = \mathbb{R}^2$, $\Omega_1 = (-\infty, 0) \times \mathbb{R}$, $\Omega_2 = (0, +\infty) \times \mathbb{R}$ and $\Gamma = \mathbb{R}$. Denoting by $\widehat{y}_i^k(x, \omega)$ the Fourier transform of $y_i^k(x, y)$ along the y-axis, (4) becomes

$$-\partial_{xx}\widehat{y}_i^k + \omega^2 \widehat{y}_i^k = 0 \text{ on } \Omega_i,\ (-1)^{i+1}\partial_x \widehat{y}_i^k = (-1)^i \left(\widehat{\lambda}^{k-1} + \rho(\widehat{y}_1^k - \widehat{y}_2^k)\right) \text{ on } \Gamma,$$

whose bounded solutions are $\widehat{y}_1^k(x,\omega) = -\frac{\widehat{\lambda}^{k-1}}{|\omega|+2\rho}\mathrm{e}^{|\omega|x}$, $\widehat{y}_2^k(x,\omega) = \frac{\widehat{\lambda}^{k-1}}{|\omega|+2\rho}\mathrm{e}^{-|\omega|x}$. The update of the multiplier (5) is then

$$\widehat{\lambda^k} = \left(1 - \frac{2\rho}{|\omega| + 2\rho}\right)\widehat{\lambda^{k-1}} = \left(\frac{|\omega|}{|\omega| + 2\rho}\right)\widehat{\lambda^{k-1}}, \tag{7}$$

which gives the convergence factor $R_1 = |\omega|/(|\omega|+2\rho)$. The method therefore is a rougher, not a smoother: it converges well for low frequencies, ω small, but stagnates for large frequencies, since $R_1 \to 1$ when $\omega \to \infty$.

2. Two multipliers and a virtual control: We consider now a different augmented Lagrangian associated to (3),

$$\mathcal{L}_\rho^{(2)}(y_i, q; \lambda_i) = \sum_{i=1}^{2} \int_{\Omega_i} \frac{1}{2}|\nabla y_i|^2 - f\, y_i \, dx + \int_\Gamma \lambda_i(y_i - q) + \frac{\rho}{2}|y_i - q|^2 \, d\sigma,$$

where we use 2 multipliers and a virtual control q for the continuity constraint $y_1|_\Gamma = y_2|_\Gamma$. Given λ^{k-1}, the first step of Algorithm 1 for $\mathcal{L}_\rho^{(2)}$ is

$$\begin{cases} -\Delta y_i^k = f \text{ in } \Omega_i,\ y_i^k|_{\partial\Omega_i \setminus \Gamma} = 0,\ \partial_{n_i} y_i^k = -\left(\lambda_i^{k-1} + \rho(y_i^k - q^k)\right) \text{ on } \Gamma, \\ 2\rho q^k = \displaystyle\sum_{i=1}^{2}\left(\rho y_i^k + \lambda_i^{k-1}\right) \text{ on } \Gamma. \end{cases} \tag{8}$$

This shows that $\partial_{n_1} y_1^k + \partial_{n_2} y_2^k = 2\rho q^k - \sum_{i=1}^2 \left(\lambda_i^{k-1} + \rho y_i^k\right) = 0$, so that the Neumann traces still match at each iteration k like in FETI. Substituting the relation in (8) for q^k on Γ into the relation for $\partial_{n_i} y_i^k$ in (8) gives

$$\partial_{n_i} y_i^k = \frac{1}{2}\left(\lambda_j^{k-1} - \lambda_i^{k-1} + \rho(y_j^k - y_i^k)\right). \tag{9}$$

The update of the multipliers from step 2 of Algorithm 1 is then

$$\lambda_i^k = \lambda_i^{k-1} + \rho\left(y_i^k - q^k\right), \tag{10}$$

and we therefore have $\partial_{n_1} y_1^k = -\lambda_1^k$ and $\partial_{n_2} y_2^k = -\lambda_2^k$. This also holds at iteration $k-1$, and substituting $\lambda_1^{k-1} = -\partial_{n_1} y_1^{k-1}$ and $\lambda_2^{k-1} = -\partial_{n_2} y_2^{k-1}$ into (9), we get on Γ the two transmission conditions

$$\begin{aligned} \partial_{n_1} y_1^k + \frac{1}{2}\rho y_1^k &= \partial_{n_1} \frac{y_1^{k-1} + y_2^{k-1}}{2} + \frac{1}{2}\rho y_2^k, \\ \partial_{n_2} y_2^k + \frac{1}{2}\rho y_2^k &= \partial_{n_2} \frac{y_1^{k-1} + y_2^{k-1}}{2} + \frac{1}{2}\rho y_1^k. \end{aligned}$$

This is again not a decoupled DD iteration because of the last terms on the right hand side at iteration k, and in addition now the normal derivatives on the right hand side are also averaged, this is no known DD iteration. We thus have again for this approach that DtO≠OtD.

Performing a Fourier analysis as in the first case, we obtain

$$\begin{cases} -\partial_{xx}\widehat{y}_i^k + \omega^2 \widehat{y}_i^k = 0 \text{ in } \Omega_i, \\ \quad (-1)^{i+1}\partial_x \widehat{y}_i^k = -\left(\widehat{\lambda}_i^{k-1} + \rho(\widehat{y}_i^k - \widehat{q}^k)\right) \text{ on } \Gamma, \\ \qquad\qquad 2\rho\widehat{q}^k = \sum_{i=1}^{2} \left(\rho\widehat{y}_i^k + \widehat{\lambda}_i^{k-1}\right) \text{ on } \Gamma. \end{cases} \tag{11}$$

The bounded solution to (11) are therefore

$$\widehat{y}_1^k = \frac{-\widehat{\lambda}_1^{k-1} + \widehat{\lambda}_2^{k-1}}{2|\omega| + 2\rho} \mathrm{e}^{|\omega| x}, \; \widehat{y}_2^k = \frac{\widehat{\lambda}_1^{k-1} - \widehat{\lambda}_2^{k-1}}{2|\omega| + 2\rho} \mathrm{e}^{-|\omega| x}, \; \widehat{q}^k = \frac{\widehat{\lambda}_1^{k-1} + \widehat{\lambda}_2^{k-1}}{2\rho}.$$

The update (10) in the algorithm then becomes

$$\widehat{\lambda}_1^k = \frac{|\omega|}{2|\omega| + 2\rho}\left(\widehat{\lambda}_1^{k-1} - \widehat{\lambda}_2^{k-1}\right), \; \widehat{\lambda}_2^k = -\frac{|\omega|}{2|\omega| + 2\rho}\left(\widehat{\lambda}_1^{k-1} - \widehat{\lambda}_2^{k-1}\right),$$

leading to the convergence factor $R_2 := \frac{|\omega|}{|\omega| + \rho} = R_1(\omega; \rho/2)$.

3 An optimization-based DDM satisfying DtO=OtD

We now use a fractional step algorithm from Glowinski [10, 5] and Le Tallec [14] to compute the saddle point of $\mathcal{L}_\rho^{(2)}$:

Algorithm 2 : Fractional step method of Glowinski (1990,2000) and Le Tallec (1995)

1. Given λ_i^k and q^{k-1}, compute y_i^k such that $\partial_{y_i}\mathcal{L}_\rho^{(2)}(y_i, q^{k-1}; \lambda_i^k) = 0$.
2. Update the multipliers as
$$\lambda_i^{k+1/2} = \lambda_i^k + \rho\partial_{\lambda_i}\mathcal{L}_\rho^{(2)}(y_i^k, q^{k-1}; \lambda_i^k).$$
3. Compute q^k satisfying $\partial_q\mathcal{L}_\rho^{(2)}(y_i^k, q^k, \lambda_i^{k+1/2}) = 0$.
4. Update the multipliers as
$$\lambda_i^{k+1} = \lambda_i^{k+1/2} + \rho\partial_{\lambda_i}\mathcal{L}_\rho^{(2)}(y_i^k, q^k; \lambda_i^{k+1/2}).$$

Applying Algorithm 2 to our model problem gives the iterates

$$\left\{\begin{aligned} -\Delta y_i^k &= f \text{ in } \Omega_i,\ y_i^k|_{\partial\Omega_i\setminus\Gamma} = 0,\ \partial_{n_i} y_i^k = -\left(\lambda_i^{k-1} + \rho(y_i^k - q^{k-1})\right) \text{ on } \Gamma,\\ \lambda_i^{k+1/2} &= \lambda_i^k + \rho(y_i^k - q^{k-1}),\\ 2\rho q^k &= \sum_{i=1}^{2}\left(\rho y_i^k + \lambda_i^{k+1/2}\right) \text{ on } \Gamma,\\ \lambda_i^{k+1} &= \lambda_i^{k+1/2} + \rho(y_i^k - q^k). \end{aligned}\right. \tag{12}$$

Substituting q^k and λ_i^k as above gives the transmission conditions

$$\partial_{n_1} y_1^{k+1} + \rho y_1^{k+1} = \partial_{n_1} y_2^k + \rho y_2^k, \quad \partial_{n_2} y_2^{k+1} + \rho y_2^{k+1} = \partial_{n_2} y_1^k + \rho y_1^k,$$

and thus the iteration (12) is the non-overlapping Schwarz method with Robin transmission conditions of P.L. Lions, DtO=OtD!

We perform again a Fourier analysis on (12) and get

$$\left\{\begin{aligned} \widehat{y}_1^k &= \tfrac{\rho\widehat{q}^{k-1} - \widehat{\lambda}_1^k}{|\omega|+\rho}\mathrm{e}^{|\omega|x},\ \widehat{y}_2^k = \tfrac{\rho\widehat{q}^{k-1} - \widehat{\lambda}_2^k}{|\omega|+\rho}\mathrm{e}^{-|\omega|x},\\ \widehat{\lambda}_i^{k+1/2} &= \tfrac{-\rho\widehat{q}^{k-1} + \widehat{\lambda}_i^k}{|\omega|+\rho},\\ \widehat{q}^k &= (\widehat{\lambda}_1^k + \widehat{\lambda}_2^k - 2\rho\widehat{q}^{k-1})\tfrac{|\omega|-\rho}{2|\omega|+2\rho},\\ \widehat{\lambda}_1^{k+1} &= \tfrac{|\omega|-\rho}{2|\omega|+2\rho}\left(\widehat{\lambda}_1^k - \widehat{\lambda}_2^k\right),\ \widehat{\lambda}_2^{k+1} = -\tfrac{|\omega|-\rho}{2|\omega|+2\rho}\left(\widehat{\lambda}_1^k - \widehat{\lambda}_2^k\right), \end{aligned}\right. \tag{13}$$

from which we obtain that the convergence factor is $R_3 := \frac{||\omega|-\rho|}{|\omega|+\rho}$.

We plot in Figure 1 the three convergence factors $R_1(\omega;\rho)$, $R_2(\omega;\rho)$ and $R_3(\omega;\rho)$. We see that we can use ρ to optimize the convergence factor for $w \in W$ where $W \subset (0,+\infty)$ is a bounded interval of frequencies. More precisely, we can control the slope of R_1, R_2 at $\omega = 0$ for the iterations (4)-(5) and (8)-(10) and, for (12), we can optimize R_3 on any bounded interval W. For the first two methods and for $w \in W$, the convergence can be made as fast as we wish by taking ρ large enough, but at the price of computing coupled subdomain solutions!

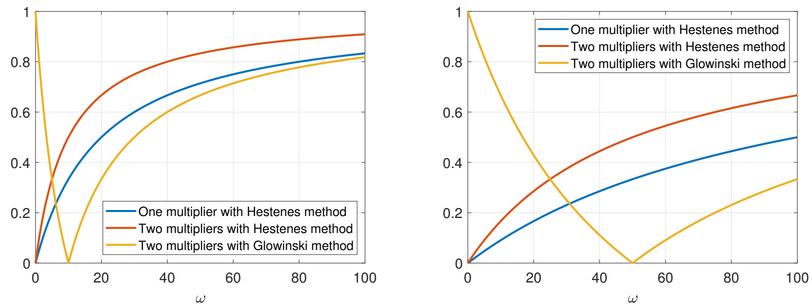

Fig. 1 Fourier convergence factors R_1, R_2 and R_3 for $\omega \in [0, 100]$. Left: $\rho = 10$. Right: $\rho = 50$.

4 Concluding remarks

We presented 3 optimization-based DDMs for the Poisson equation written as a minimization problem following the Decompose-then-Optimize (DtO) approach, and showed that the first two lead to unknown domain decomposition methods for the first order stationarity conditions, DtO≠OtD (Optimize-then-Decompose), while the last one leads to the non-overlapping parallel Schwarz method of Lions and we have DtO=OtD.

The major difference comes from the augmented Lagrangian algorithm used in DtO to derive these methods: in Algorithm 2, the virtual control is updated independently of the solution, which leads to a fully decoupled algorithm with Robin transmission conditions, while with Algorithm 1, the Robin conditions and thus subdomain solves do not decouple across the iterations.

We can now also answer the DtO=OtD question for the method introduced in [3] for a linear quadratic optimization problem based on Algorithm 1 using one multiplier for the continuity constraint. Fitting it into our DtO framework leads to the augmented Lagrangian

$$\mathcal{L}_\rho^{(\mathrm{opt})}(g, u_i; \lambda) = \frac{1}{2}\sum_{i=1}^{2} \|y_i - y_\mathrm{t}\|^2_{L^2(\Omega_i)} + \alpha\|u_i\|^2_{L^2(\Omega_i)} + \int_\Gamma \lambda(y_1 - y_2) + \frac{\rho}{2}\int_\Gamma (y_1 - y_2)^2,$$

where the $y_i := y_i(g, u_i)$ satisfy $-\Delta y_i = f + u_i$ in Ω_i, $y_i|_{\partial\Omega_i\setminus\Gamma} = 0$ and $\partial_{n_i} y_i = (-1)^{i+1} g$ on Γ. Applying Algorithm 1 to $\mathcal{L}_\rho^{(\mathrm{opt})}$ yields an iteration involving both the primal variables y_i and adjoint variables p_i which satisfy $-\Delta p_i^k = -(y_i^k - y_\mathrm{t})$ in Ω_i, $p_i^k|_{\partial\Omega_i\setminus\Gamma} = 0$ and $\partial_{n_i} p_i^k = (-1)^i \left(\lambda^{k-1} + \rho(y_1^k - y_2^k)\right)$ on Γ. Since $\partial_{n_i} p_i^{k-1} = (-1)^i \lambda^{k-1}$, each iteration of this algorithm satisfies on Γ the transmission conditions $\partial_{n_1} p_1^k + \rho y_1^k = \partial_{n_1} p_2^{k-1} + \rho y_2^k$ and $\partial_{n_2} p_2^k + \rho y_2^k = \partial_{n_2} p_1^{k-1} + \rho y_1^k$ that are similar to (6), and hence for the DtO method from [3] we also have DtO≠OtD.

References

1. Ciaramella, G., Halpern, L., Mechelli, L.: Convergence analysis and optimization of a Robin Schwarz waveform relaxation method for periodic parabolic optimal control problems. Tech. rep., MOX Lab, Politecnico di Milano (2022)
2. Ciaramella, G., Kwok, F., Müller, G.: Nonlinear optimized Schwarz preconditioner for elliptic optimal control problems. In: Domain Decomposition Methods in Science and Engineering XXVI. Springer (2020)
3. Cocquet, P.H., Gander, M.J., Vieira, A.: Decompose-then-optimize: a new approach to design domain decomposition methods for optimal control problems. Numerical Algorithms **in print** (2025)
4. Farhat, C., Roux, F.X.: 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)
5. Fortin, M., Glowinski, R.: Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems. Elsevier (2000)
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, pp. 207–216. Springer (2016)
7. Gander, M.J., Kwok, F., Mandal, B.C.: Convergence of substructuring methods for elliptic optimal control problems. In: Domain Decomposition Methods in Science and Engineering XXIV 24, pp. 291–300. Springer (2018)
8. Gander, M.J., Leugering, G.: Nonoverlapping domain decomposition of parabolic optimal control problems revisited: decompose-then-optimize versus optimize-then-decompose. In: Submitted (2024)
9. Gander, M.J., Lu, L.D.: Dirichlet-Neumann and Neumann-Neumann methods for elliptic control problems. In: Domain decomposition methods in science and engineering XXVII, pp. 207–216. Springer (2023)
10. 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, pp. 224–231. SIAM Philadelphia (1990)
11. Hestenes, M.R.: Multiplier and gradient methods. Journal of optimization theory and applications **4**(5), 303–320 (1969)
12. Hinze, M., Pinnau, R., Ulbrich, M., Ulbrich, S.: Optimization with PDE constraints, vol. 23. Springer Science & Business Media (2008)
13. Lagnese, J.E., Leugering, G.: Domain in Decomposition Methods in Optimal Control of Partial Differential Equations, vol. 148. Springer Science & Business Media (2004)
14. Le Tallec, P., Sassi, T.: Domain decomposition with nonmatching grids: augmented lagrangian approach. Mathematics of Computation **64**(212), 1367–1396 (1995)
15. Leugering, G.: Domain decomposition of optimal control problems for dynamic networks of elastic strings. Computational Optimization and Applications **16**, 5–27 (2000)
16. Leugering, G.: Nonoverlapping domain decomposition for instantaneous optimal control of friction dominated flow in a gas-network. Tech. rep. (2020). Preprint
17. Powell, M.J.: A method for nonlinear constraints in minimization problems. Optimization pp. 283–298 (1969)
18. Vieira, A., Cocquet, P.H.: Optimized Schwarz method for coupled direct-adjoint problems applied to parameter identification in advection-diffusion equation. In: Accepted for publication in proceeding of Domain Decomposition Methods in Science and Engineering XXVII. Springer (2023)