Domain Decomposition Methods in Gradient Descent Prodcedures for Optimal Control Problems

IX: Partial differential equations, optimal design and numerics
Benasque, 2022

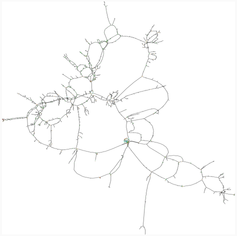
Lukas Wolff, Falk Hante, Daniël Veldman, Enrique Zuazua

26.08.2022

FAU Erlangen-Nürnberg, Subproject C07 of TRR154 of the DFG

Introduction

Motivation



Overarching goal: Optimization of gas network operation

I.e.: Optimal control of nonlinear PDE's on large domains/networks:

$$\min_{u \in \mathcal{K}} J(y(u), u)$$

with

- target functional J (operational cost, distance to target state)
- control u (compressor stations, valves)
- state y (pressure, density, velocity)
- set of admissible controls $\mathcal{K}$

Optimal control of gas flow on networks

Let $(\mathcal{E},\mathcal{V})$ a network with boundary nodes $\mathcal{V}_b$, interior nodes $\mathcal{V}_0$ and compressor nodes $\mathcal{V}_c$. Our goal is to solve

$$\min_{u\in U} J(y,u) := \int_{\mathcal{E}\times[0,T]} (y-y_d)^2 + \frac{\delta}{2}\|u\|_U^2 \tag{1}$$

such that

$$\begin{cases} \partial_t(\varphi(y_e)) - \partial_x(\varphi(\partial_x y_e)) & = 0 \quad \text{in } \{e\}\times(0,T), \quad e\in\mathcal{E} \\ \nu_e\varphi(\partial_x y) + \gamma_v\varphi(y) & = g \quad \text{in } \{v\}\times(0,T) \quad v\in\mathcal{V}_b,\ e\in\mathcal{E}(v) \\ y_e & = y_{e'} \quad \text{in } \{v\}\times(0,T), \quad v\in\mathcal{V}_0,\ e,e'\in\mathcal{E}(v) \\ \sum_{e\in\mathcal{E}(v)}\nu_e\varphi(\partial_x y_e) & = 0 \quad \text{in } \{v\}\times(0,T), \quad v\in\mathcal{V}_0\setminus\mathcal{V}_c \\ \nu_e\varphi(\partial_x y_e) & = \gamma_v(\varphi(y_{e'}) - \varphi(y_e) + u_e) \\ & \qquad \text{in } \{v\}\times(0,T),\ v\in\mathcal{V}_c,\ e,e'\in\mathcal{E}(v) \\ y_e(0,\cdot) & = y_{e,0} \quad \text{on } e\in\mathcal{E} \end{cases} \tag{2}$$

with nonlinearity $\varphi : s\mapsto |s|^{-\frac{1}{2}}s$. Our project (C07 of TRR154) concerns gradient descent procedures for this kind of problem:

- Does the equation allow for an adjoint state?
- Can we decrease numerical cost of the GD by using domain decomposition?

Optimal control of parabolic p-Laplacian equations

For $\alpha < 2$ consider problems of the form

$$\min_{u \in U} J(y,u) = \frac{1}{2}\int_0^T \int_\Omega |y - y_d|^2 + \frac{\lambda}{2}\int_0^T \int_{\partial\Omega} u^2 \tag{3a}$$

$$\text{s.t.} \begin{cases} \partial_t y - \Delta_\alpha y & = 0 \quad \text{in } (0,T)\times\Omega, \\ |\nabla y|^{\alpha-2}\dfrac{\partial y}{\partial \nu} + \gamma y & = \gamma u \quad \text{on } (0,T)\times\partial\Omega, \\ y(x,0) & = y_0 \quad \text{in } \Omega \end{cases} \tag{3b}$$

for some fixed $y_0 \in L^2(\Omega)$, $\gamma, \lambda > 0$, $y_d \in L^2(\Omega)$ and a control $u \in U \subset L^{\frac{\alpha}{\alpha-1}}(0,T,L^2(\partial\Omega))$. The formal adjoint p can become singular:

$$-\partial_t p - \nabla \cdot \left(\varphi'(\nabla y)\nabla p\right) = 0$$

Open problem: Assuming only $\nabla y \neq 0$ a.e., is the solution operator of (3) differentiable? [cf. Fernandez-Casas, 1995]

Theorem (LW, Zuazua, 2022)
Assuming $0 < C_1 \leq |\nabla y| \leq C_2 < \infty$, the solution operator of (3) is Gâteaux differentiable and an adjoint state exists.

Domain Decomposition for Gradient Descent

Motivation: Gradient Descent Procedures for Optimal Control

For simplicity, consider

$$\min_{u \in L^2(\Omega)} J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Omega)} \tag{4}$$

$$\text{with } \begin{cases} -\Delta y &= f + u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

Optimality System: Optimal state $\bar{y}$ and its adjoint $\bar{p}$ satisfy

$$\begin{cases} -\Delta \bar{y} &= f - \frac{1}{\lambda}\bar{p} \quad \text{in } \Omega \\ \bar{y} &= 0 \quad \text{on } \partial\Omega \end{cases}, \qquad \begin{cases} -\Delta \bar{p} &= \bar{y} - y_d \quad \text{in } \Omega \\ \bar{p} &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{5}$$

Gradient Descent: Find $\bar{y}$ by iteratively computing

$$\begin{cases} -\Delta y^{(n)} &= f + u^{(n)} \quad \text{in } \Omega \\ y^{(n)} &= 0 \quad \text{on } \partial\Omega \end{cases}, \qquad \begin{cases} -\Delta p^{(n)} &= y^{(n)} - y_d \quad \text{in } \Omega \\ p^{(n)} &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{6}$$

$$u^{(n+1)} = u^{(n)} - \eta\left(p^{(n)} + \lambda u^{(n)}\right) \tag{7}$$

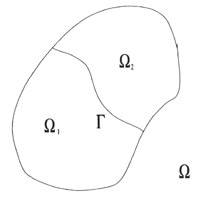
Main challenge: Repeated computation of forward and adjoint system is numerically expensive! Particularly for very large domains.

Domain Decomposition Methods

Consider the Poisson problem

$$\begin{cases} -\Delta y &= f \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{8}$$

on a domain $\Omega \subset \mathbb{R}^d$. Fix a decomposition $\Omega = \overline{\Omega}_1 \cup \overline{\Omega}_2$ with $\Omega_1 \cap \Omega_2 = \emptyset$ and $\Gamma := \overline{\Omega}_1 \cap \overline{\Omega}_2$.



Picture taken from [LL04]

Then, the iterative procedure defined by

$$\begin{cases} -\Delta y_i^n &= f|_{\Omega_i} \quad \text{in } \Omega_i \\ y_i^n &= 0 \quad \text{on } \partial\Omega_i \setminus \Gamma \\ \partial_{\nu_i} y_i^n + y_i^n &= -\partial_{\nu_j} y_j^{n-1} + y_j^{n-1} \quad \text{on } \Gamma, j \neq i \end{cases} \tag{9}$$

converges towards y for $n \to \infty$ [Lio90]. Note: parallelizable!

Domain Decomposition and Gradient Descent

How can DDM be used in the context of gradient descent?

1. **Performing DDM in each iteration of GD** (see [CY11]):
   - Outer Iteration: Adjoint-based gradient descent
   - Inner Iteration: Solve forward and adjoint system by DDM
2. **Decomposing the optimal control problem:**
   - Outer Iteration: DDM for optimality systems (see [BD96], [LL04])
   - Inner Iteration: Solve decomposed optimality systems by gradient descent

Idea: Can a "diagonal" approach be performed?

Proposed Procedure

Iteratively for $n = 1, 2, \ldots$ do:

1. Compute the state equation in all Ω_i in parallel:

$$\begin{cases} -\Delta y_i^n & = f|_{\Omega_i} + u^n|_{\Omega_i} \quad \text{in } \Omega_i \\ y_i^n & = 0 \quad \text{on } \partial\Omega_i \setminus \Gamma_{ij} \\ \partial_\nu y_i^n + y_i^n & = -\partial_{\nu_j} y_j^{n-1} + y_j^{n-1} \quad \text{on } \Gamma, j \neq i \end{cases} \tag{10}$$

2. Compute the adjoint equation in all Ω_i in parallel:

$$\begin{cases} -\Delta p_i^n & = y_i^n - y_d|_{\Omega_i} \quad \text{in } \Omega_i \\ p_i^n & = 0 \quad \text{on } \partial\Omega_i \setminus \Gamma \\ \partial_\nu p_i^n + p_i^n & = -\partial_{\nu_j} p_j^{n-1} + p_j^{n-1} \quad \text{on } \Gamma_{ij}, j \neq i \end{cases} \tag{11}$$

3. Update the controls in Ω_i based on p_i^n.

Questions:

- Can convergence be ensured?
- Can a numerical advantage be achieved?

Results

Numerical Experiments

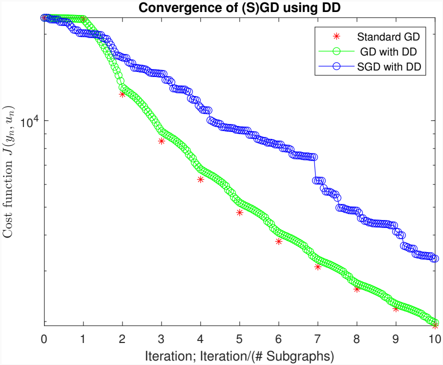


Figure 1: $|\mathcal{V}| = 422$, $|\mathcal{E}| = 1143$, Number of Subgraphs $M = 20$

Descent Procedure	Time [s]
Standard GD	97
GD with DD	68
SGD with DD	72

Convergence Analysis

Theorem (Hante, LW, Veldman, Zuazua, 2022)
We can chose the stepsize $\eta > 0$ *sufficiently small such that for the optimal control problem* (4) *on networks and 1D domains the proposed procedure converges towards the optimal control* $\bar{u}$.

Idea of the proof: For small stepsizes, the change in the control in each step is small enough to not disturb the convergence of the domain decomposition method.

Framework of the proof

$\mathcal{H}$,$\mathcal{U}$ Hilbert spaces, $\mathcal{X}$, $\mathcal{Y}$ Banach spaces,$A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $B \in \mathcal{L}(\mathcal{U}, \mathcal{Y})$, $C \in \mathcal{L}(\mathcal{X}, \mathcal{H})$. Consider the optimal control problem

$$\begin{aligned} &\min_{u \in U} J(y,u) = \tfrac{1}{2}\|Cy - z_d\|^2_{\mathcal{H}} + \tfrac{1}{2}\|u\|^2_{\mathcal{U}} \\ &\text{s.t. } Ay = Bu + f, \end{aligned} \tag{12}$$

for some given $z_d \in \mathcal{H}$ and $f \in \mathcal{Y}$. Note that $Ay = f$ can be solved by the iteration

$$My_{n+1} = Ny_n + f \tag{13}$$

for $A = M - N$ and $\rho(M^{-1}N) < 1$. Consider the following descent procedure: For $n = 0, 1, 2, \ldots$ compute

$$\begin{cases} My_{n+1} &= Ny_n + Bu_n + f, \\ \tilde{M}p_{n+1} &= \tilde{N}p_n + C^*(Cy_{n+1} - z_d), \\ u_{n+1} &\leftarrow u_n - \eta(B^*p_{n+1} + u_n). \end{cases} \tag{14}$$

Proposition

If $\rho(M^{-1}N) < 1$ and $\rho(\tilde{M}^{-1}\tilde{N}) < 1$, there exists $\eta_0 > 0$ such that for all stepsizes $0 < \eta < \eta_0$ the algorithm (14) converges for all initial guesses.

How to understand DDM in this framework?

Let $Ay = f$ denote the poisson equation on Ω. We can then decompose

$$A = \begin{bmatrix} M_1 & N_1 \\ N_2 & M_2 \end{bmatrix} = \underbrace{\begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix}}_{=:M} - \underbrace{\begin{bmatrix} 0 & -N_1 \\ -N_2 & 0 \end{bmatrix}}_{=:N}$$

where $M_i y_i = N_i y_j + f$ corresponds to

$$\begin{cases} -\Delta y_i & = f|_{\Omega_i} \quad \text{in } \Omega_i \\ y_i & = 0 \quad \text{on } \partial\Omega_i \setminus \Gamma \\ \partial_{\nu_i} y_i + y_i & = -\partial_{\nu_j} y_j + y_j \quad \text{on } \Gamma, j \neq i \end{cases}$$

where we decomposed $\Omega = \Omega_1 \cup \Omega_2$. This can be made precise!

Sketch of proof of the proposition

Proof.
The algorithm (14) can be written in matrix form as

$$\underbrace{\begin{bmatrix} M & 0 & 0 \\ -C^*C & \tilde{M} & 0 \\ 0 & \eta B^* & I_{\mathcal{U}} \end{bmatrix}}_{\mathcal{M}_2(\eta)} \begin{bmatrix} y_{k+1} \\ p_{k+1} \\ u_{k+1} \end{bmatrix} - \underbrace{\begin{bmatrix} N & 0 & B \\ 0 & \tilde{N} & 0 \\ 0 & 0 & I_{\mathcal{U}} - \eta I_{\mathcal{U}} \end{bmatrix}}_{\mathcal{N}_2(\eta)} \begin{bmatrix} y_k \\ p_k \\ u_k \end{bmatrix} = \begin{bmatrix} f \\ -C^* z_d \\ 0 \end{bmatrix}.$$

Establishing the convergence comes down to showing that $\rho(\mathcal{M}_2^{-1}(\eta)\mathcal{N}_2(\eta)) < 1$. Using $\rho(M^{-1}N) < 1$ and $\rho(\tilde{M}^{-1}\tilde{N}) < 1$, this can be done for $\eta > 0$ small enough.

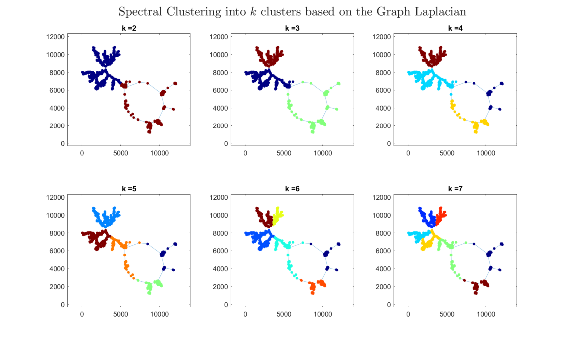
□

Outlook

Outlook

Future goals:

- Extend our convergence results, particularly to instationary problems
- Investigate possible advantages of choosing subdomains stochastically
- Analyze the influence of the decomposition method on the convergence properties, particularly on networks



Thank you for your attention!

Are there any Questions?

References

[BD96] J.-D. Benamou and B. Desprès. *A Domain Decomposition Method for the Helmholtz Equation and Related Optimal Control Problems*. Research Report RR-2791. Projet IDENT. INRIA, 1996. URL: https://hal.inria.fr/inria-00073899.

[CY11] H. Chang and D. Yang. A Schwarz domain decomposition method with gradient projection for optimal control governed by elliptic partial differential equations. In: *Journal of computational and applied mathematics* 235.17 (2011), pp. 5078–5094.

[Lio90] P.-L. Lions. “On the Schwarz alternating method. III: a variant for nonoverlapping subdomains”. In: *Third international symposium on domain decomposition methods for partial differential equations*. Vol. 6. SIAM Philadelphia, PA. 1990, pp. 202–223.

[LL04] J. E. Lagnese and G. Leugering. *Domain Decomposition Methods in Optimal Control of Partial Differential Equations*. Basel: Birkhäuser Basel, 2004. ISBN: 978-3-0348-9610-8. DOI: 10.1007/978-3-0348-7885-2.