A Non-Conforming Dual Approach for Adaptive Trust-Region Reduced Basis Approximation of PDE-Constrained Parameter Optimization

Seminar Advanced Numerical Methods and Applications

Peter Oehme

22.07.2022

Download Handout and Slides

Alternatively:
github.com/peoe/sem-nc-dual-trrb-pdeopt

Outline

1. General Overview
2. Theoretical Framework
   - PDE Constrained Optimization
   - Trust-Region and BFGS Method
3. Optimality Systems and *a posteriori* Error Estimates
4. Adaptive Trust-Region Reduced Basis Algorithm and Convergence
   - The TR-RB Algorithm
   - Convergence of the TR-RB Algorithm
5. Numerical Results

Parametrized Partial Differential Equations

RB Problem

Rectangular parameter space $\mathcal{P} := \{\mu \in \mathbb{R}^d \mid (\mu_a)_i \leq \mu_i \leq (\mu_b)_i, i = 1, \ldots, d\}$
Find a solution $u \in V$ for

$$a_\mu(u, v) = l_\mu(v) \qquad \forall v \in V$$

- a_μ: continuous, coercive, bilinear, **parameter separable**, and
- l_μ: continuous, linear, **parameter separable**.

Optimization over Parameter Domain

Parameter Optimization

$$\min_{\mu} \mathcal{J}(u, \mu)$$

$$\mathcal{J}(u, \mu) := \Theta(\mu) + j_\mu(u) + k_\mu(u, u)$$

- k_μ: continuous, symmetric, bilinear,
- j_μ: continuous, linear, and
- Θ: arbitrary parameter function.

Adaptive Approach

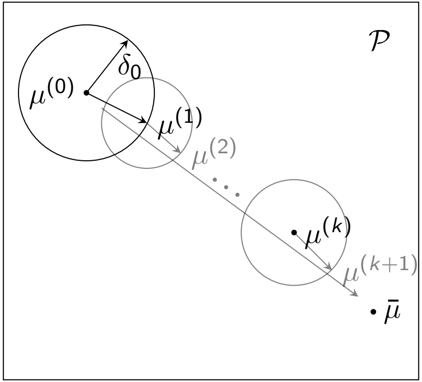


Figure: Local adaptation (TR) on the domain, figure courtesy of Tim Keil

Adaptive Approach

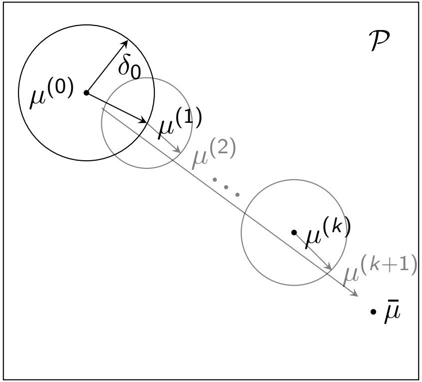


Figure: Local adaptation (TR) on the domain, figure courtesy of Tim Keil

while *not converged* **do**
  Compute locally optimal solution;
  if *some condition* **then**
    Continue to next optimization step;
  else
    Change some parameter of the problem;
    Optimize again;

Adaptive Approach

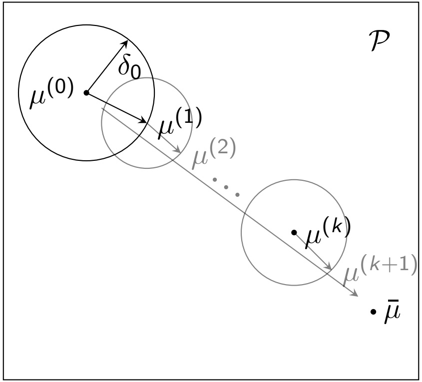


Figure: Local adaptation (TR) on the domain, figure courtesy of Tim Keil

while *not converged* **do**
  Compute locally optimal solution;
  if *some condition* **then**
    Continue with refinement;
  else if *some other condition* **then**
    Repeat optimization step with smaller domain;
  else if *some third condition* **then**
    Refine and possibly repeat optimization step;

Adaptive Approach

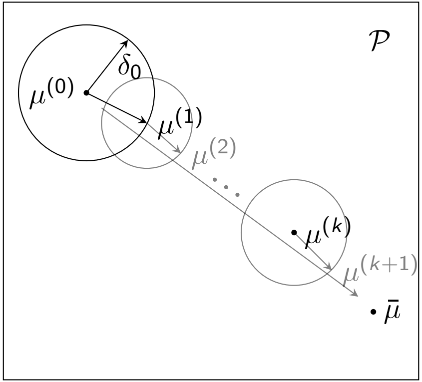


Figure: Local adaptation (TR) on the domain, figure courtesy of Tim Keil

Benefits of the Adaptive Approach

We only adapt locally.

We **do not** construct a reduced basis for the entire domain $\mathcal{P}$!

PDE Constrained Optimization

Motivation

Parameter dependence in PDEs:

- material properties
- velocities
- heat sources/sinks
- ...

How to determine optimal parameters?

PDE Constrained Optimization

Problem Description, c.f. [Hin+09]

We consider

$$\min_{\mu \in \mathcal{P}} \mathcal{J}(u, \mu) \qquad s.t.$$
$$e(u, \mu) = 0,$$

where $e_\mu(u, v)$ encodes the PDE constraint

$$e(u, \mu) := l_\mu(v) - a_\mu(u, v) \qquad \forall v \in V \text{ or } V_N$$

Notes on Differentiability

Directional Derivative, c.f. [Hin+09]

$$dF(x)[h] := \lim_{t \to 0} \frac{F(x+th) - F(x)}{t} \in Y$$

Gâteaux Derivative, c.f. [Hin+09]

$$dF(x) \in \mathcal{L}(X, Y)$$

Fréchet Derivative, c.f. [Hin+09]

$$\|F(x+th) - F(x) - dF(x)[h]\|_Y = o(\|h\|_X), \|h\|_X \to 0$$

Notes on Differentiability

Bilinear Form

Chain rule:

$$
\begin{aligned}
d_\mu a_\mu(u_\mu, v_\mu) \cdot \nu &= \partial_\mu a_\mu(u_\mu, v_\mu) \cdot \nu + \partial_u a_\mu(u_\mu, v_\mu)[d_\nu u_\mu] + \partial_v a_\mu(u_\mu, v_\mu)[d_\nu v_\mu] \\
&= \partial_\mu a_\mu(u_\mu, v_\mu) \cdot \nu + a_\mu(d_\nu u_\mu, v_\mu) + a_\mu(u_\mu, d_\nu v_\mu)
\end{aligned}
$$

$d_\nu u_\mu, d_\nu v_\mu$ are also called **sensitivities**.

Trust-Region Method

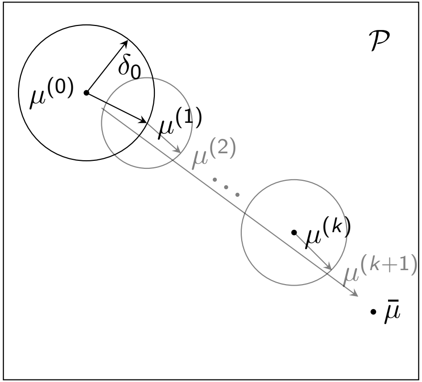


Figure: Local adaptation (TR) on the domain, figure courtesy of Tim Keil

Model function, c.f. [Kel99]

Idea: Replace with simpler function $m^{(k)}$ modelling $\mathcal{J}$ around some point $x^{(k)}$.

Trust-Region Method

Requirements, c.f. [Qia+17]

- $|\mathcal{J}(u_\mu, \mu) - \mathcal{J}_N(\mu)| \leq \Delta_{\mathcal{J}_N}(\mu)$,
- $\|\nabla_\mu \mathcal{J}(u_\mu, \mu) - \nabla_\mu \mathcal{J}_N(\mu)\|_2 \leq \Delta_{\nabla \mathcal{J}_N}(\mu)$, and
- $\mathcal{J}_N^{(k+1)}(\mu^{(k+1)}) \leq \mathcal{J}_N^{(k)}(\mu^{(k,0)})$

The last is equivalent to

$$\mathcal{J}_N^{(k)}(\mu^{(k+1)}) + \Delta_{\mathcal{J}_N^{(k)}}(\mu^{(k+1)}) < \mathcal{J}_N^{(k)}(\mu^{(k)})$$

BFGS Method

Newton Method, c.f. [NW06]

$$x^{(k+1)} := x^{(k)} + \kappa^{(k)} d^{(k)},$$

$$d^{(k)} := -\mathcal{H}(x^{(k)})^{-1} \nabla \mathcal{J}(x^{(k)})$$

BFGS Method

Newton Method, c.f. [NW06]

$$x^{(k+1)} := x^{(k)} + \kappa^{(k)} d^{(k)},$$

$$d^{(k)} := -\mathcal{H}(x^{(k)})^{-1} \nabla \mathcal{J}(x^{(k)})$$

Quasi Newton Method, c.f. [Kel99; NW06]

$$x^{(k+1)} := x^{(k)} + \kappa^{(k)} d^{(k)},$$

$$d^{(k)} := -\mathcal{H}^{(k)^{-1}} \nabla \mathcal{J}(x^{(k)})$$

BFGS Method

BFGS Update Formula

$$\mathcal{H}^{(k+1)} := \mathcal{H}^{(k)} + \frac{(y^{(k)} - \mathcal{H}^{(k)} s^{(k)})(y^{(k)} - \mathcal{H}^{(k)} s^{(k)})^T}{(y^{(k)} - \mathcal{H}^{(k)} s^{(k)})^T s^{(k)}},$$

$$s^{(k)} := x^{(k+1)} - x^{(k)},$$

$$y^{(k+1)} := \nabla \mathcal{J}(x^{(k+1)}) - \nabla \mathcal{J}(x^{(k)})$$

Primal and Dual Equations

Primal Residual

For $u, v \in V, \mu \in \mathcal{P}$ we define

$$Res_\mu^{pr}(u)[v] := l_\mu(v) - a_\mu(u, v).$$

Dual Residual

For $u, v, p \in V, \mu \in \mathcal{P}$ we define

$$Res_\mu^{du}(u, p)[v] := j_\mu(v) + 2k_\mu(v, u) - a_\mu(v, p).$$

Optimality Conditions

Lagrangian: $\mathcal{L}(u, \mu, p) := \mathcal{J}(u, \mu) - Res_\mu^{pr}(u)[p]$

First Order Necessary Conditions, c.f. [Kei+21]

For any locally optimal pair (u, μ) there exists an optimal dual variable $p \in V$ such that the following conditions are satisfied

$$\begin{aligned} a_\mu(u, v) &= l_\mu(v) && \forall v \in V \\ a_\mu(u, p) &= \partial_u \mathcal{J}(u, \mu)[v] && \forall v \in V \\ 0 &\leq \nabla_\mu \mathcal{L}(u, \mu, p) \cdot (\nu - \mu) && \forall \nu \in \mathcal{P}. \end{aligned}$$

Optimality Conditions — Standard Approach

Reduced Functional

For $\mu \in \mathcal{P}$ we define

$$J_N(\mu) := \mathcal{J}(u_{N,\mu}, \mu),$$

where $u_{N,\mu}$ defines the reduced solution for μ.

Then the gradient for all $p \in V_N^{du}$ is

$$(\nabla_\mu J_N(\mu))_i = \partial_u \mathcal{L}(u_{N,\mu}, \mu, p)[d_{\mu_i} u_{N,\mu}] + \partial_{\mu_i} \mathcal{L}(u_{N,\mu}, \mu, p) \quad \forall p \in V_N^{du}.$$

Optimality Conditions — NCD-Corrected Approach

NCD-Corrected Functional, c.f. [Kei+21]

For $u_{N,\mu} \in V_N^{pr}$, $p_{N,\mu} \in V_N^{du}$ we define

$$\mathcal{J}_N(\mu) := \mathcal{L}(u_{N,\mu}, \mu, p_{N,\mu}) = J_N(\mu) + Res_\mu^{pr}(u_{N,\mu})[p_{N,\mu}]$$

Optimality Conditions — NCD-Corrected Approach

NCD-Corrected Gradient, Standard Approach, c.f. [Kei+21]

For $v_{N,\mu} \in V_N^{pr}, p_{N,\mu} \in V_N^{du}$ we get

$$\begin{aligned}(\nabla_\mu \mathcal{J}_N(\mu))_i &= \partial_{\mu_i} \mathcal{J}(u_{N,\mu}, \mu) \\ &\quad + \partial_{\mu_i} Res_\mu^{pr}(u_{N,\mu})[p_{N,\mu} + w_{N,\mu}] - \partial_{\mu_i} Res_\mu^{du}(u_{N,\mu}, p_{N,\mu})[z_{N,\mu}],\end{aligned}$$

where $z_{N,\mu} \in V_N^{du}, w_{N,\mu} \in V_N^{pr}$ satisfy

$$\begin{aligned} a_\mu(z_{N,\mu}, q) &= -Res_\mu^{pr}(u_{N,\mu})[q] && \forall q \in V_N^{du}, \text{ and} \\ a_\mu(v, w_{N,\mu}) &= Res_\mu^{du}(u_{N,\mu}, p_{N,\mu})[v] - 2k_\mu(z_{N,\mu}, v) && \forall v \in V_N^{pr}. \end{aligned}$$

a posteriori Error Estimates

Model Reduction Error, c.f. [Qia+17; Kei+21]

For some $\mu \in \mathcal{P}$ we consider the solutions $u_\mu \in V$, $u_{N,\mu} \in V_N^{pr}$, $p_\mu \in V$, $p_{N,\mu} \in V_N^{du}$. Then it holds that

$$\left\| u_\mu - u_{N,\mu} \right\| \leq \Delta_{pr}(\mu) := {\alpha_\mu}^{-1} \left\| Res_\mu^{pr}(u_{N,\mu}) \right\|, \text{ and}$$

$$\left\| p_\mu - p_{N,\mu} \right\| \leq \Delta_{du}(\mu) := {\alpha_\mu}^{-1} \left(2\gamma_{k_\mu} \Delta_{pr}(\mu) + \left\| Res_\mu^{du}(u_{N,\mu}, p_{N,\mu}) \right\| \right).$$

a posteriori Error Estimates

$$\|u_\mu - u_{N,\mu}\| \leq \Delta_{pr}(\mu), \qquad \|p_\mu - p_{N,\mu}\| \leq \Delta_{du}(\mu)$$

Output Model Reduction Error, c.f. [Qia+17; Kei+21]

For $\mu \in \mathcal{P}, u_\mu \in V, u_{N,\mu} \in V_N^{pr}, p_\mu \in V, p_{N,\mu} \in V_N^{du}$ we have

$$\begin{aligned} |\mathcal{J}(u_\mu, \mu) - J_N(\mu)| \leq \Delta_{J_N}(\mu) &:= \Delta_{pr}(\mu) \left\| Res_\mu^{du}(u_{N,\mu}, p_{N,\mu}) \right\| \\ &\quad + \Delta_{pr}(\mu)^2 \gamma_{k_\mu} + \left| Res_\mu^{pr}(u_{N,\mu})[p_{N,\mu}] \right|, \text{ and} \\ |\mathcal{J}(u_\mu, \mu) - \mathcal{J}_N(\mu)| \leq \Delta_{\mathcal{J}_N}(\mu) &:= \Delta_{pr}(\mu) \left\| Res_\mu^{du}(u_{N,\mu}, p_{N,\mu}) \right\| + \Delta_{pr}(\mu)^2 \gamma_{k_\mu}. \end{aligned}$$

a posteriori Error Estimates

$$|\mathcal{J}(u_\mu, \mu) - \mathcal{J}_N(\mu)| \leq \Delta_{\mathcal{J}_N}(\mu)$$

Output Gradient Model Reduction Error, c.f. [Qia+17; Kei+21]

For $\mu \in \mathcal{P}, u_\mu \in V, u_{N,\mu} \in V_N^{pr}, p_\mu \in V, p_{N,\mu} \in V_N^{du}$ we have

$$\left\|\nabla_\mu \mathcal{J}(u_\mu, \mu) - \nabla_\mu \mathcal{J}_N(\mu)\right\|_2 \leq \Delta_{\nabla \mathcal{J}_N}(\mu) := \left\|\Delta^*_{\nabla \mathcal{J}_N}(\mu)\right\|_2 .$$

TR-RB Algorithm

Choice of Model Function

We choose

$$m^{(k)}(\eta) := \mathcal{J}_N^{(k)}(\mu^{(k)} + \eta),$$

where $\eta \in \mathcal{P}$.

Note: $\mathcal{J}_N \neq \mathcal{J}_N^{(k)}$! These are based upon different spaces $V_N^{pr,k}, V_N^{du,k}$.

TR-RB Algorithm

---

while *not converged* **do**
    Compute locally optimal solution;
    if *some condition* **then**
        Continue with refinement;
    else if *some other condition* **then**
        Repeat optimization step with smaller domain;
    else if *some third condition* **then**
        Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**

Compute locally optimal solution;

if $\mathcal{J}_N^{(k)}(\mu^{(k+1)}) + \Delta_{\mathcal{J}_N^{(k)}}(\mu^{(k+1)}) < \mathcal{J}_N^{(k)}(\mu^{(k)})$ **then**

Continue with refinement;

else if *some other condition* **then**

Repeat optimization step with smaller domain;

else if *some third condition* **then**

Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**
    Compute locally optimal solution;
    if *output decrease sufficiently small* **then**
        Continue with refinement;
    else if *some other condition* **then**
        Repeat optimization step with smaller domain;
    else if *some third condition* **then**
        Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**
  Compute locally optimal solution;
  if *output decrease sufficiently small* **then**
    Continue with refinement;
  else if $\mathcal{J}_N^{(k)}(\mu^{(k+1)}) + \Delta_{\mathcal{J}_N^{(k)}}(\mu^{(k+1)}) > \mathcal{J}_N^{(k)}(\mu^{(k)})$ **then**
    Repeat optimization step with smaller domain;
  else if *some third condition* **then**
    Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**
    Compute locally optimal solution;
    if *output decrease sufficiently small* **then**
        Continue with refinement;
    else if *output decrease not small enough* **then**
        Repeat optimization step with smaller domain;
    else if *some third condition* **then**
        Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**
  Compute locally optimal solution;
  if *output decrease sufficiently small* **then**
    Continue with refinement;
  else if *output decrease not small enough* **then**
    Repeat optimization step with smaller domain;
  else
    Refine and possibly repeat optimization step;

TR-RB Algorithm

while *not converged* **do**
  Compute locally optimal solution;
  if *output decrease sufficiently small* **then**
    Refine and potentially enlarge trust-radius;
  else if *output decrease not small enough* **then**
    Repeat optimization step with smaller domain;
  else
    Refine and check whether to repeat optimization (shrinking, enlarging, or keeping of trust-radius depending on outcome);

TR-RB Algorithm

Inner Loop (BFGS), c.f. [Qia+17; Kei+21]

The inner (BFGS) step is

$$\mu^{(k,l+1)} := \mu^{(k,l)}(j) := \mathbb{P}(\mu^{(k,l)} + \kappa^j d^{(k,l)}),$$

where

$$\mathbb{P}(\mu)_i := \begin{cases} (\mu_a)_i, & (\mu)_i \leq (\mu_a)_i, \\ (\mu)_i, & (\mu_a)_i \leq (\mu)_i \leq (\mu_b)_i, \\ (\mu_b)_i, & (\mu_b)_i \leq (\mu)_i \end{cases}.$$

Finally,

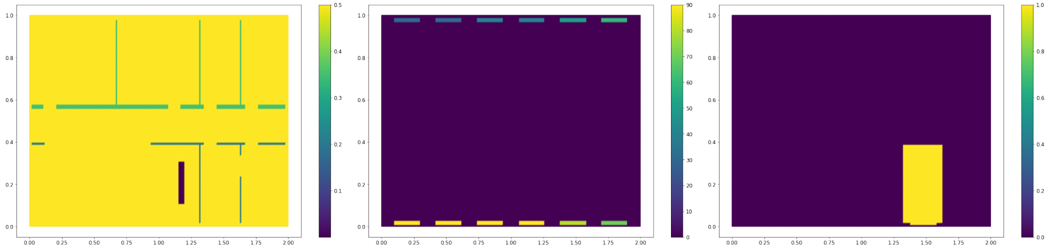
$$\mu^{(k+1)} := \mu^{(k,L)}.$$

Overview on Convergence

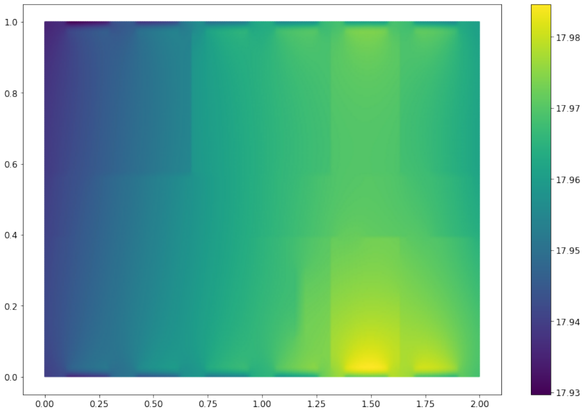
Proper **convergence is not shown**, just

All accumulation points of the sequence of parameters are approximate first order critical points.

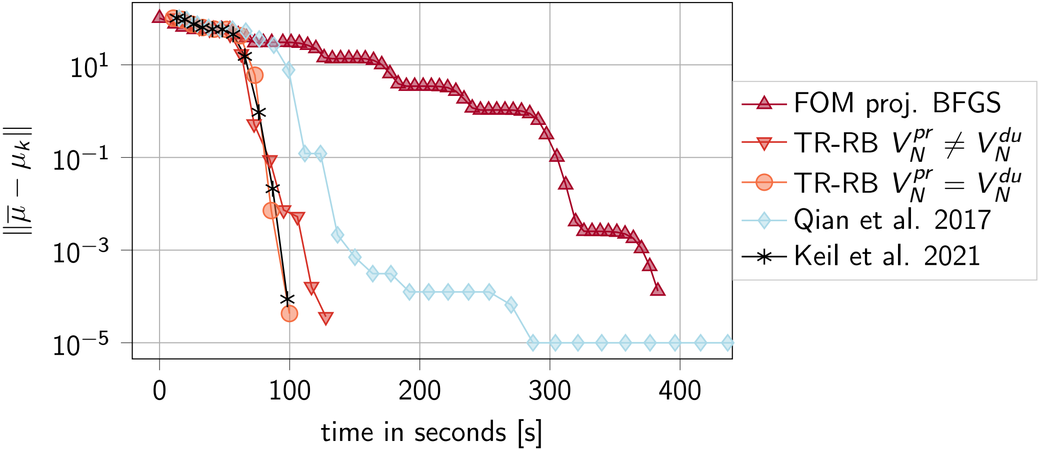
Numerical Results, c.f. github.com/TiKeil



Numerical Results, c.f. github.com/TiKeil



Numerical Results, c.f. [Kei+21], data courtesy of Tim Keil



[Hin+09] Michael Hinze, Rene Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE Constraints*. Springer Netherlands, 2009. DOI: 10.1007/978-1-4020-8839-1.

[Kei+21] Tim Keil, Luca Mechelli, Mario Ohlberger, Felix Schindler, and Stefan Volkwein. "A non-conforming dual approach for adaptive Trust-Region reduced basis approximation of PDE-constrained parameter optimization". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 55 (3 May 2021), pp. 1239–1269. ISSN: 28047214. DOI: 10.1051/m2an/2021019.

[Kel99] C. T. Kelley. *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics, Jan. 1999. DOI: 10.1137/1.9781611970920.

[NW06] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer New York, 2006. ISBN: 978-0-387-30303-1. DOI: 10.1007/978-0-387-40065-5.

[Qia+17] Elizabeth Qian, Martin Grepl, Karen Veroy, and Karen Willcox. "A Certified Trust Region Reduced Basis Approach to PDE-Constrained Optimization". In: *SIAM Journal on Scientific Computing* 39 (5 Jan. 2017), S434–S460. ISSN: 1064-8275. DOI: 10.1137/16m1081981.

Thank you for your attention!

Questions/Discussion