Universität
Münster

Reduced Order Surrogate Models for PDE-Constrained Optimization & Inverse Problems

Mario Ohlberger

with B Haasdonk, M Kartmann, H Kleikamp, B Klein, S Rave,
F Schindler, T Schuster, S. Volkwein, T. Wenzel

Symposium of Yvon Maday

Collège de France, Paris, June 21-24, 2026

living.knowledge

MM
Mathematics
Münster
Cluster of Excellence

Outline

Motivation: Multi-fidelity learning for PDE constrained Optimization

A certified and adaptive RB-ML-ROM surrogate approach

PDE-constrained Optimization: A Trust Region RB-ML-ROM approach

Inverse Parameter Identification: RB-TR-IRGNM

Outline

Motivation: Multi-fidelity learning for PDE constrained Optimization

A certified and adaptive RB-ML-ROM surrogate approach

PDE-constrained Optimization: A Trust Region RB-ML-ROM approach

Inverse Parameter Identification: RB-TR-IRGNM

Motivation: Multi-fidelity learning for PDE constrained Optimization

We consider PDE-constrained optimization problems of the form

$$\min_{\mu \in \mathcal{P}_{\mathrm{ad}}} J(\mu, u_\mu), \text{ such that } e(u_\mu; \mu) = 0 \text{ in } W',$$

where $J : \mathcal{P} \times V \to \mathbb{R}$ denotes the objective functional and $e(\cdot; \mu) : V \to W'$ a residual encoding the parameter dependent primal state equation (PDE).

Motivation: Multi-fidelity learning for PDE constrained Optimization

We consider PDE-constrained optimization problems of the form

$$\min_{\mu \in \mathcal{P}_{\mathrm{ad}}} J(\mu, u_\mu), \text{ such that } e(u_\mu; \mu) = 0 \text{ in } W',$$

where $J : \mathcal{P} \times V \to \mathbb{R}$ denotes the objective functional and $e(\cdot; \mu) : V \to W'$ a residual encoding the parameter dependent primal state equation (PDE).

We consider descent methods, i.e.,

$$\mu_{k+1} := \mu_k + t_k d_k(\mu_k, u_{\mu_k}, p_{\mu_k}, \dots),$$

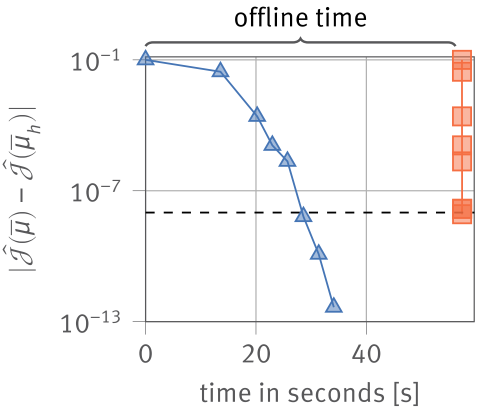
where t_k and d_k are suitably chosen step sizes and descent directions, respectively.

Classical RB Approach: Timings for PDE Constrained Optimization

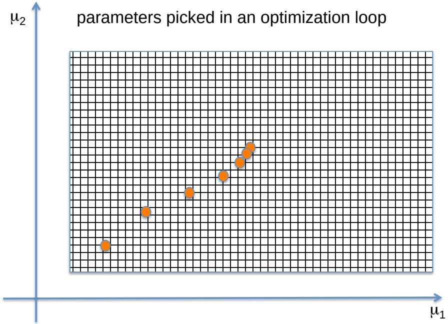
Full Order Model (FOM)

vs.

Classical Reduced Order Model (ROM)

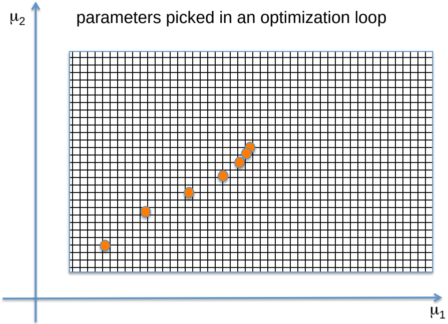


Idea: Learning of Surrogate Models on the Fly



$\Longrightarrow$ Surrogate model only needed along a path in some points.

Idea: Learning of Surrogate Models on the Fly



$\Longrightarrow$ Surrogate model only needed along a path in some points.

Multi-fidelity approach: Replace primal and dual sate solutions to compute descent directions by a hierarchy of models, e.g. FOM, RB, and ML models that are learned on-the-fly.

Outline

Motivation: Multi-fidelity learning for PDE constrained Optimization

A certified and adaptive RB-ML-ROM surrogate approach

PDE-constrained Optimization: A Trust Region RB-ML-ROM approach

Inverse Parameter Identification: RB-TR-IRGNM

A certified and adaptive RB-ML-ROM surrogate approach

[Haasdonk, Kleikamp, Ohlberger, Schindler, Wenzel. SISC 2023]

Parametrized parabolic PDE in weak form: Find $u(\mu) \in L^2(0,T;V)$ s.t.

$$\begin{aligned} \langle \partial_t u(\mu), v\rangle + a\big(u(\mu), v; \mu\big) &= l(v;\mu), && \text{for all } v \in V,\\ u(0;\mu) &= u_0(\mu), && \text{for } u_0(\mu) \in V, \end{aligned}$$

with coercive bilinear form a, linear functional l and initial condition u_0.

Application:

- Many-query scenarios, such as parameter optimization or optimal control problems

Major obstacle:

- Large number of time steps, also in the reduced method!

A certified and adaptive RB-ML-ROM surrogate approach

[Haasdonk, Kleikamp, Ohlberger, Schindler, Wenzel. SISC 2023]

Parametrized parabolic PDE in weak form: Find $u(\mu) \in L^2(0, T; V)$ s.t.

$$\begin{aligned} \langle \partial_t u(\mu), v \rangle + a\big(u(\mu), v; \mu\big) &= l(v; \mu), && \text{for all } v \in V, \\ u(0; \mu) &= u_0(\mu), && \text{for } u_0(\mu) \in V, \end{aligned}$$

with coercive bilinear form a, linear functional l and initial condition u_0.

Application:

- Many-query scenarios, such as parameter optimization or optimal control problems

Major obstacle:

- Large number of time steps, also in the reduced method!

Idea: Combined Machine Learning and Model Order Reduction Approach!

Machine learning surrogate models

- Let V_N denote a reduced space of dimension n with basis $\{\varphi_i\}_{i=0}^n$, i.e. $u_N(\mu; t_k) = \sum_{i=1}^n U_i(\mu; t_k)\varphi_i$.

Machine learning surrogate models

- Let V_N denote a reduced space of dimension n with basis $\{\varphi_i\}_{i=0}^{n}$, i.e. $u_N(\mu; t_k) = \sum_{i=1}^{n} U_i(\mu; t_k)\varphi_i$.

Idea: Learn map from parameter $\mu \in \mathcal{P}$ to coefficient vector $\underline{U(\mu; t_k)} \in \mathbb{R}^N$ using machine learning.

Machine learning surrogate models

- Let V_N denote a reduced space of dimension n with basis $\{\varphi_i\}_{i=0}^n$, i.e. $u_N(\mu;t_k)=\sum_{i=1}^n U_i(\mu;t_k)\varphi_i$.

Idea: Learn map from parameter $\mu \in \mathcal{P}$ to coefficient vector $\underline{U(\mu;t_k)} \in \mathbb{R}^N$ using machine learning.

Two different approaches:

1. Time-"vectorized":
   Train $\Phi\colon \mathbb{R}^p \to \mathbb{R}^{K\times N}, \quad \mu \to [U_i(\mu;t_k)]_{k=1}^K,$ with $p := \dim \mathcal{P}$, $K = \#$time steps.
   $\Longrightarrow$ predict the ROM coefficients for all time steps at once.
2. "Random-access"-in-time:
   Train $\Phi\colon \mathbb{R}^{p+1} \to \mathbb{R}^N, \quad (\mu,t) \to U_i(\mu;t)$
   $\Longrightarrow$ predict the ROM coefficients at any point in time.

Machine learning surrogate models

- Let V_N denote a reduced space of dimension n with basis $\{\varphi_i\}_{i=0}^{n}$, i.e. $u_N(\mu; t_k)=\sum_{i=1}^{n} U_i(\mu; t_k)\varphi_i$.

Idea: Learn map from parameter $\mu \in \mathcal{P}$ to coefficient vector $\underline{U(\mu; t_k)} \in \mathbb{R}^N$ using machine learning.

Two different approaches:

1. Time-"vectorized":
   Train $\Phi\colon \mathbb{R}^p \to \mathbb{R}^{K\times N}, \quad \mu \to [U_i(\mu; t_k)]_{k=1}^{K},$ with $p := \dim \mathcal{P}$, $K = \#$time steps.
   $\Longrightarrow$ predict the ROM coefficients for all time steps at once.
2. "Random-access"-in-time:
   Train $\Phi\colon \mathbb{R}^{p+1} \to \mathbb{R}^N, \quad (\mu, t) \to U_i(\mu; t)$
   $\Longrightarrow$ predict the ROM coefficients at any point in time.

- Denote DoF-vector $\underline{U_{\mathrm{ML}}(\mu; t_k)} := \Phi(\mu)_k$ (or $\Phi(\mu, t^k)) \in \mathbb{R}^N$ and corresponding solution $u_{\mathrm{ML}}(\mu; t_k) \in V_N$.

Related Approaches

[Hesthaven and Ubbiali, Non-intrusive ROMs using neural networks, 2018]

[Kutyniok, Petersen, Raslan, Schneider, Theoretical analysis of DNNs and parametric PDEs, 2022]

[Fresca and Manzoni, POD-DL-ROM, 2022]

....

[Brivio, Fresca, Franco, Manzoni., Error estimates for POD-DL-ROMs, 2024]

[Chellappa, Feng, Benner, Error estimation via data-enhanced error closure, 2024]

and many more ....

Comparison of the models

Full-order/high-fidelity model (FOM)

Pros:

- arbitrarily accurate solutions

Cons:

- very slow when considering fine discretizations and many time steps

Reduced order model (ROM)

Pros:

- faster than FOM
- error estimator

Cons:

- slow iterative time stepping
- solve dense linear system in each time step

Machine learning model (MLM)

Pros:

- faster than ROM
- parallel evaluation for different time instances

Cons:

- no certification available (so far)

Parabolic RB a posteriori error estimator

Offline-online-decomposable error estimate for difference between FOM-solution $u_h(\mu) \in V_h$ and ROM-solution $u_{\mathrm{RB}}(\mu) \in V_N$

$$\|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_{L^2(0,T;V_h)} \leq E\big(u_{\mathrm{RB}}(\mu);\mu\big)$$

$$:= \underbrace{\alpha(\mu)}_{\substack{\text{coercivity}\\ \text{constant}}}{}^{-1} \left(\sum_{k=1}^{K-1} \Delta t \| \underbrace{R_k\big(u_{\mathrm{RB}}(t_k;\mu);\mu\big)}_{\substack{\text{residuum in the}\\ k\text{-th time step}}} \|_{V_h'}^2 \right)^{1/2}$$

Parabolic RB a posteriori error estimator

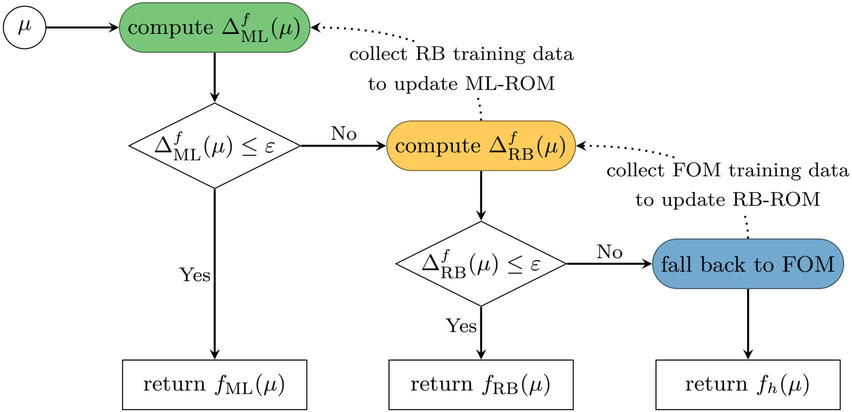
Offline-online-decomposable error estimate for difference between FOM-solution $u_h(\mu) \in V_h$ and ROM-solution $u_{\mathrm{RB}}(\mu) \in V_N$

$$
\begin{aligned}
\|u_h(\mu) - u_{\mathrm{RB}}(\mu)\|_{L^2(0,T;V_h)} &\leq E\left(u_{\mathrm{RB}}(\mu);\mu\right) \\
&:= \underbrace{\alpha(\mu)}_{\substack{\text{coercivity}\\ \text{constant}}}{}^{-1} \left(\sum_{k=1}^{K-1} \Delta t \| \underbrace{R_k\left(u_{\mathrm{RB}}(t_k;\mu);\mu\right)}_{\substack{\text{residuum in the}\\ k\text{-th time step}}} \|_{V_h'}^2 \right)^{1/2}
\end{aligned}
$$

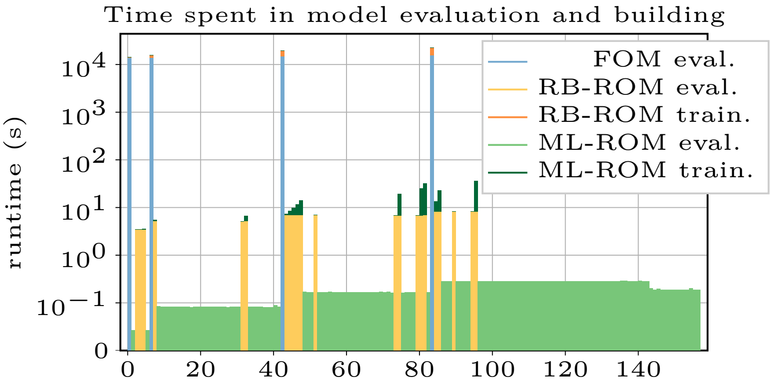
Idea: Apply error estimator E to machine learning prediction $u_{\mathrm{ML}}(\mu)$!

$\Longrightarrow$ A posteriori certification of machine learning results!

Adaptive multi-fidelity model



Experiment – PDE-solves with VKOGA as ML model



Outline

Motivation: Multi-fidelity learning for PDE constrained Optimization

A certified and adaptive RB-ML-ROM surrogate approach

PDE-constrained Optimization: A Trust Region RB-ML-ROM approach

Inverse Parameter Identification: RB-TR-IRGNM

Trust Region RB-ML-ROM for parabolic PDE Constrained Optimization

[Klein, Ohlberger, ACOM 2026+]

- Objective functional $\mathcal{J} : V \times \mathcal{P} \to \mathbb{R}$

$$\mathcal{J}(u_\mu;\mu) := \Delta t \sum_{k=1}^{K} \|u_\mu^k - g_{\text{ref}}^k\|_V^2 + \lambda \mathcal{R}(\mu),$$

PDE-Constrained Optimization

Find the minimizer of the reduced objective functional

$$\bar{\mu} := \operatorname*{arg\,min}_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu) := \mathcal{J}(u_\mu, \mu),$$

subject to $u_\mu := (u_\mu^k)_{k \in \{0,\dots,K\}} \in V^{K+1}$ being the solution of a time-discretized parabolic primal PDE

$$\frac{(u_\mu^k - u_\mu^{k-1}, v)_{L^2(\Omega)}}{\Delta t} + a(u_\mu^k, v; \mu) = b(t^k) f(v;\mu)$$

$\forall v \in V$, $K \in \{1, \dots K\}$ and $u_\mu^0 = 0$.

Trust Region RB-ML-ROM for parabolic PDE Constrained Optimization

[Klein, Ohlberger, ACOM 2026+]

- Objective functional $\mathcal{J} : V \times \mathcal{P} \to \mathbb{R}$

$$\mathcal{J}(u_\mu; \mu) := \Delta t \sum_{k=1}^{K} \|u_\mu^k - g_{\text{ref}}^k\|_V^2 + \lambda \mathcal{R}(\mu),$$

PDE-Constrained Optimization

Find the minimizer of the reduced objective functional

$$\bar{\mu} := \arg\min_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu) := \mathcal{J}(u_\mu, \mu),$$

subject to $u_\mu := (u_\mu^k)_{k \in \{0,\ldots,K\}} \in V^{K+1}$ being the solution of a time-discretized parabolic primal PDE

$$\frac{(u_\mu^k - u_\mu^{k-1}, v)_{L^2(\Omega)}}{\Delta t} + a(u_\mu^k, v; \mu) = b(t^k) f(v; \mu)$$

$\forall v \in V,\ K \in \{1, \ldots K\}$ and $u_\mu^0 = 0$.

Challenges for optimization:

- efficient evaluation of $\hat{\mathcal{J}}(\mu)$ and Gradient $\nabla \hat{\mathcal{J}}(\mu)$
  $\Rightarrow$ efficient evaluation of primal & dual state equations
- high dimensional parameter space $\mathcal{P}$
- multiscale/large scale context

Efficient optimization method using RB-ML-ROM approach.

Trust Region methods

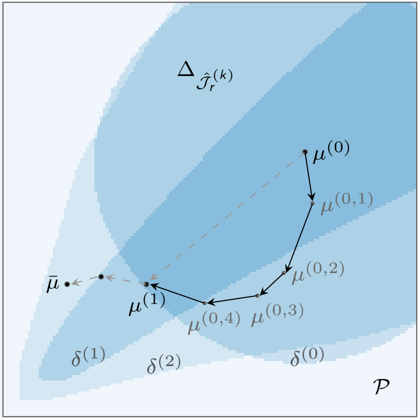
Error Aware Trust Region Optimization

Approximate $\bar{\mu} = \arg\min_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu)$

by solving optimization subproblems

$$\mu_{k+1} := \mu_k + s_k, \qquad \min_{s \in \mathcal{P}} m_k(s), \quad s.t. \frac{\Delta_{\hat{\mathcal{J}}_n^{(k)}}(\tilde{\mu})}{\hat{\mathcal{J}}_n^{(k)}(\tilde{\mu})} \leq \delta_k,$$

starting with some m_0, δ_0.



Trust Region methods

Error Aware Trust Region Optimization

Approximate $\bar{\mu} = \underset{\mu \in \mathcal{P}}{\arg\min}\, \hat{\mathcal{J}}(\mu)$

by solving optimization subproblems

$$\mu_{k+1} := \mu_k + s_k, \qquad \min_{s \in \mathcal{P}} m_k(s), \quad s.t. \frac{\Delta_{\hat{\mathcal{J}}_n^{(k)}}(\tilde{\mu})}{\hat{\mathcal{J}}_n^{(k)}(\tilde{\mu})} \leq \delta_k,$$

starting with some m_0, δ_0.

Idea:

- $m_k(s) := \hat{\mathcal{J}}_n^k(\mu_k + s)$ (RB-ML-ROM reduced $\hat{\mathcal{J}}$)
- Subproblem solver: BFGS method

Requirements:

- **model function** $m_k \approx \hat{\mathcal{J}}$, smooth and $\|\nabla m_k\| < \infty$
- **characterization** of the Trust Region
- bound $\Delta_{\hat{\mathcal{J}}_n^{(k)}}(\mu)$ on $|m_k(\mu) - \hat{\mathcal{J}}(\mu)|$ for all $\mu \in \mathcal{P}$
- solver for the optimization subproblem

New:

- Multifidelity RB-ML-ROM approach
- applied to primal and dual state equation for efficient computation of first order derivatives
- constrained optimization
- full error control and convergence

First-optimize-then-discretize

For a **descent method** (e.g. BFGS) the following quantities are required:

primal state
$$\frac{1}{\Delta t}(u_\mu^k - u_\mu^{k-1}, v)_{L^2(\Omega)} + a(u_\mu^k, v; \mu) = b(t^k) f(v; \mu), \quad u^0(\mu) = 0$$

dual state
$$\frac{1}{\Delta t}(v, p_\mu^k - p_\mu^{k+1})_{L^2(\Omega)} + a(v, p_\mu^k; \mu) = 2d(u_\mu^k, v) + l^k(v), \quad p^{K+1}(\mu) = 0$$

reduced gradient
$$\nabla_\mu \hat{\mathcal{J}}(\mu) = \nabla_\mu \mathcal{L}(u_\mu, p_\mu, \mu) \quad (\mathcal{L} \text{ from Lagrangian ansatz (all-at-once)})$$

Idea: Use RB-ML-ROM approach for an efficient evaluation of both, primal and dual states and use resulting multi-fidelity surrogate as model function in the Trust Region subproblems!

A posteriori error estimation

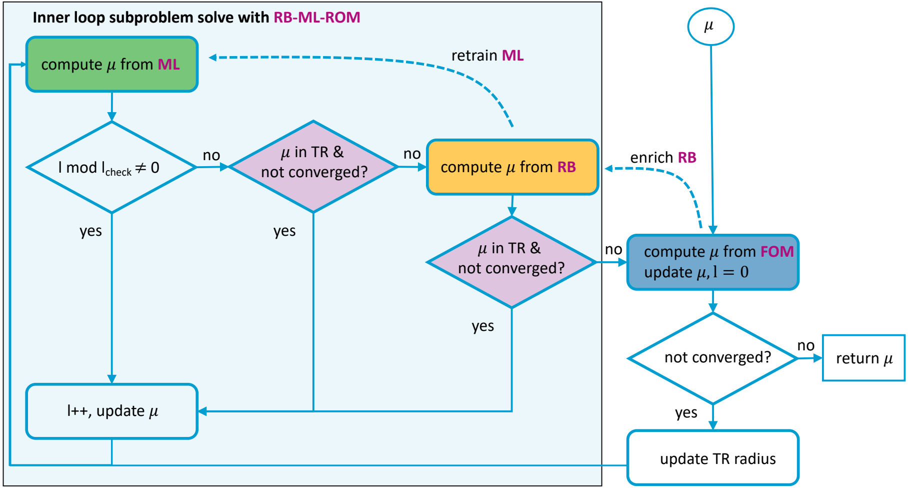
Required for error control and adaptivity

$$
\begin{aligned}
\|u_{h,\mu} - u_{n,\mu}\|_h &\lesssim \Delta^n_{\mathrm{pr}}(\mu) \\
\|p_{h,\mu} - p_{n,\mu}\|_h &\lesssim \Delta^n_{\mathrm{du}}(\mu) \\
|\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_n(\mu)| &\lesssim \Delta^n_{\hat{\mathcal{J}}}(\mu) \\
\left|\left(\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_n(\mu)\right)_i\right| &\lesssim \Delta^n_{\nabla_\mu \hat{\mathcal{J}}_i}(\mu)
\end{aligned}
$$

$\Longrightarrow$ Fully developed theory!

based on [Qian, Grepl, Veroy, Willcox. SISC 2017] &
[Keil, Mechelli, Ohlberger, Schindler, Volkwein, ESAIM M2AN, 2021]

Relaxed Trust Region RB-ML-ROM



Numerical Experiment

Model problem

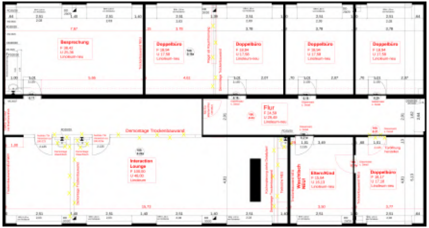
Quadratic objective functional:

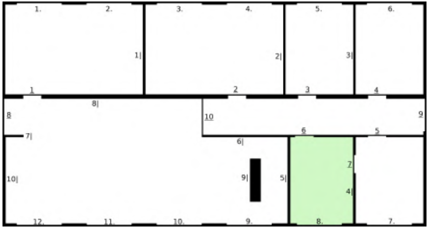
$$\mathcal{J}(u,\mu) := \frac{\sigma_d}{2}\int_D (u - u^{\mathrm{d}})^2 + \frac{1}{2}\sum_{i=1}^{M}\sigma_i(\mu_i - \mu_i^d)^2,$$

Parametrized constrained equation: Find u_μ s.t.

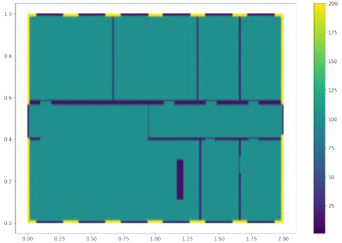
$$\begin{aligned} \partial_t u_\mu - \nabla\cdot(\kappa_\mu \nabla u_\mu) &= f_\mu && \text{in } \Omega, \\ q(\kappa_\mu \nabla u_\mu \cdot n) &= (u_{\mathrm{out}} - u_\mu) && \text{on } \partial\Omega. \end{aligned}$$

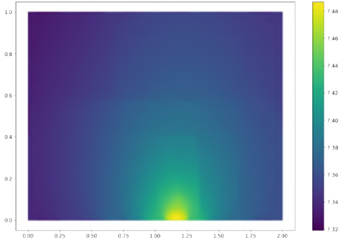
- Optimizing heat distribution in a building.
- Two-dimensional parameter optimization.
- $K = 10.000$ and $N_h = 10.151$.





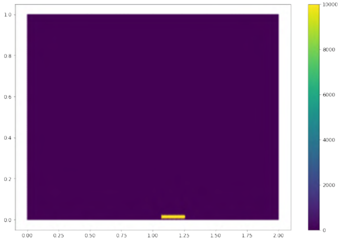
heat conductivity κ_μ

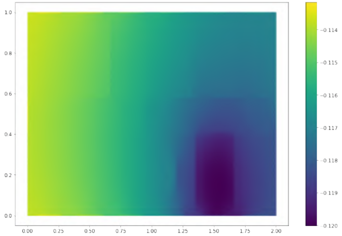




primal state solution u_μ

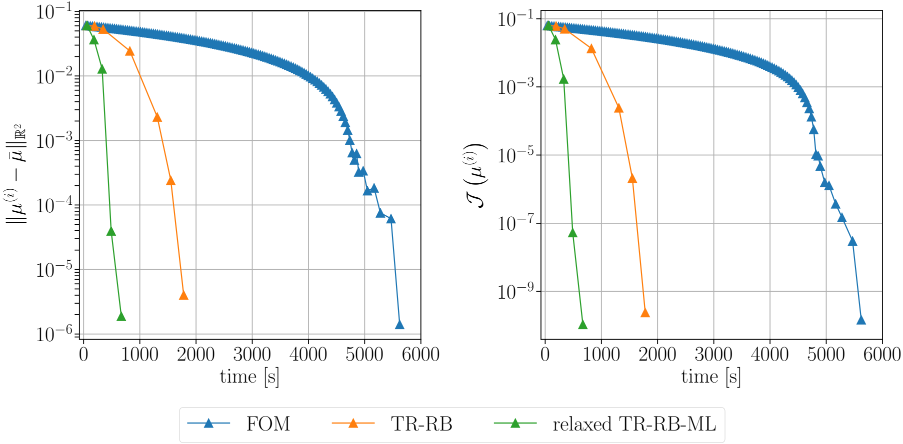
source f_μ



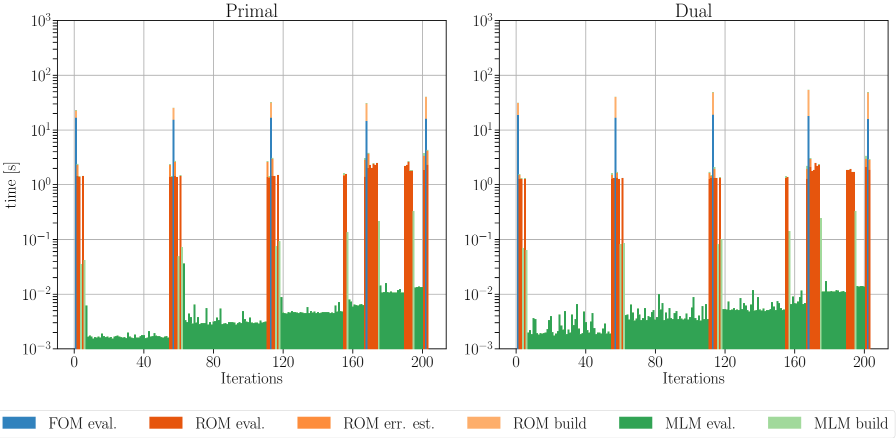


dual state solution p_μ

Error versus CPU time



CPU time per primal/dual solve versus iteration



Computational efficiency

Algorithm	time $[s]$	FOM solves	RB solves	MLM solves	Speed up
FOM	4706	121.80	0.00	0.00	–
TR-RB	1447	6.30	143.00	0.00	3.25
relaxed TR-RB-ML	557	4.30	35.10	148.90	8.45

Outline

Motivation: Multi-fidelity learning for PDE constrained Optimization

A certified and adaptive RB-ML-ROM surrogate approach

PDE-constrained Optimization: A Trust Region RB-ML-ROM approach

Inverse Parameter Identification: RB-TR-IRGNM

Reduced Basis Trust Region Methods for Inverse Parameter Identification

[Kartmann, Keil, Ohlberger, Volkwein, Kaltenbacher. Comput. Sci. Eng. 2024]
[Kartmann, Klein, Ohlberger, Volkwein, Schuster. Inverse Problems 2025]

Let Q, C, H, V be real Hilbert spaces, V is dense in H.
Let $\mathscr{Q}_{\text{ad}} \subset \mathscr{Q} \subset L^2(0,T;Q), \mathscr{C} := L^2(0,T;C)$,
$\mathscr{V} := L^2(0,T;V) \cap H^1(0,T;V')$.
For $q \in \mathscr{Q}_{\text{ad}}$, let $u = S(q) \in \mathscr{V}$ denote the solution of

$$\begin{aligned} \langle \partial_t u(t) + A(q(t))u(t), v \rangle_{V',V} &= \langle l(t), v \rangle_{V',V}, \\ u(0) &= 0 \text{ in } H. \end{aligned}$$

For $q \in \mathscr{Q}_{\text{ad}}$ given, the measurement $y \in \mathscr{C}$ is given by the forward operator

$$\mathcal{F}(q) = y, \qquad \text{with } \mathcal{F} : \mathscr{Q}_{\text{ad}} \to \mathscr{C}, \ \mathcal{F} = \mathcal{C} \circ \mathcal{S}$$

where $\mathcal{C} : \mathscr{V} \to \mathscr{C}$ is a linear bounded operator.

Reduced Basis Trust Region Methods for Inverse Parameter Identification

[Kartmann, Keil, Ohlberger, Volkwein, Kaltenbacher. Comput. Sci. Eng. 2024]
[Kartmann, Klein, Ohlberger, Volkwein, Schuster. Inverse Problems 2025]

Let Q, C, H, V be real Hilbert spaces, V is dense in H.
Let $\mathscr{Q}_{\text{ad}} \subset \mathscr{Q} \subset L^2(0,T;Q), \mathscr{C} := L^2(0,T;C)$,
$\mathscr{V} := L^2(0,T;V) \cap H^1(0,T;V')$.
For $q \in \mathscr{Q}_{\text{ad}}$, let $u = S(q) \in \mathscr{V}$ denote the solution of

$$\begin{aligned} \langle \partial_t u(t) + A(q(t))u(t), v\rangle_{V',V} &= \langle l(t), v\rangle_{V',V}, \\ u(0) &= 0 \text{ in } H. \end{aligned}$$

For $q \in \mathscr{Q}_{\text{ad}}$ given, the measurement $y \in \mathscr{C}$ is given by the forward operator

$$\mathcal{F}(q) = y, \qquad \text{with } \mathcal{F} : \mathscr{Q}_{\text{ad}} \to \mathscr{C}, \ \mathcal{F} = \mathcal{C} \circ \mathcal{S}$$

where $\mathcal{C} : \mathscr{V} \to \mathscr{C}$ is a linear bounded operator.

Inverse parameter identification

Given noisy measurements $y^\delta \in \mathcal{C}$, solve

$$\min J(q) := \frac{1}{2} \|\mathcal{F}(q) - y^\delta\|_{\mathcal{C}}^2 \quad \text{s.t.} \quad q \in \mathscr{Q}_{\text{ad}}.$$

Reduced Basis Trust Region Methods for Inverse Parameter Identification

[Kartmann, Keil, Ohlberger, Volkwein, Kaltenbacher. Comput. Sci. Eng. 2024]
[Kartmann, Klein, Ohlberger, Volkwein, Schuster. Inverse Problems 2025]

Let Q, C, H, V be real Hilbert spaces, V is dense in H.
Let $\mathscr{Q}_{\text{ad}} \subset \mathscr{Q} \subset L^2(0,T;Q), \mathscr{C} := L^2(0,T;C)$,
$\mathscr{V} := L^2(0,T;V) \cap H^1(0,T;V')$.
For $q \in \mathscr{Q}_{\text{ad}}$, let $u = S(q) \in \mathscr{V}$ denote the solution of

$$\langle \partial_t u(t) + A(q(t))u(t), v \rangle_{V',V} = \langle l(t), v \rangle_{V',V},$$
$$u(0) = 0 \text{ in } H.$$

For $q \in \mathscr{Q}_{\text{ad}}$ given, the measurement $y \in \mathscr{C}$ is given by the forward operator

$$\mathcal{F}(q) = y, \qquad \text{with } \mathcal{F} : \mathscr{Q}_{\text{ad}} \to \mathscr{C}, \ \mathcal{F} = \mathcal{C} \circ \mathcal{S}$$

where $\mathcal{C} : \mathscr{V} \to \mathscr{C}$ is a linear bounded operator.

Inverse parameter identification

Given noisy measurements $y^\delta \in \mathcal{C}$, solve

$$\min J(q) := \frac{1}{2} \|\mathcal{F}(q) - y^\delta\|_{\mathcal{C}}^2 \quad \text{s.t.} \quad q \in \mathscr{Q}_{\text{ad}}.$$

Challenges:

- problem is ill posed ($\mathcal{F}$ not continuously invertible)
  $\Rightarrow$ Iteratively regularized Gauß-Newton method
- infinite/very high dimensional parameter space $\mathscr{Q}_{\text{ad}}$
  $\Rightarrow$ need for parameter space reduction

Reduced Basis Trust Region Methods for Inverse Parameter Identification

[Kartmann, Keil, Ohlberger, Volkwein, Kaltenbacher. Comput. Sci. Eng. 2024]
[Kartmann, Klein, Ohlberger, Volkwein, Schuster. Inverse Problems 2025]

Let Q, C, H, V be real Hilbert spaces, V is dense in H.
Let $\mathscr{Q}_{\text{ad}} \subset \mathscr{Q} \subset L^2(0,T;Q), \mathscr{C} := L^2(0,T;C)$,
$\mathscr{V} := L^2(0,T;V) \cap H^1(0,T;V')$.
For $q \in \mathscr{Q}_{\text{ad}}$, let $u = S(q) \in \mathscr{V}$ denote the solution of

$$\langle \partial_t u(t) + A(q(t))u(t), v \rangle_{V',V} = \langle l(t), v \rangle_{V',V},$$
$$u(0) = 0 \text{ in } H.$$

For $q \in \mathscr{Q}_{\text{ad}}$ given, the measurement $y \in \mathscr{C}$ is given by the forward operator

$$\mathcal{F}(q) = y, \qquad \text{with } \mathcal{F} : \mathscr{Q}_{\text{ad}} \to \mathscr{C}, \ \mathcal{F} = \mathcal{C} \circ \mathcal{S}$$

where $\mathcal{C} : \mathscr{V} \to \mathscr{C}$ is a linear bounded operator.

Inverse parameter identification

Given noisy measurements $y^\delta \in \mathcal{C}$, solve

$$\min J(q) := \frac{1}{2} \|\mathcal{F}(q) - y^\delta\|_{\mathcal{C}}^2 \quad \text{s.t.} \quad q \in \mathscr{Q}_{\text{ad}}.$$

Challenges:

- problem is ill posed ($\mathcal{F}$ not continuously invertible)
  $\Rightarrow$ Iteratively regularized Gauß-Newton method
- infinite/very high dimensional parameter space $\mathscr{Q}_{\text{ad}}$
  $\Rightarrow$ need for parameter space reduction

Efficient solution method using combined state and parameter MOR techniques.

Iterately regularized Gauß-Newton method [Kaltenbacher, Neubauer, Scherzer 2008]

Let $q^k \in \mathscr{Q}_{\text{ad}}$ be a given iterate. The IRGNM update scheme results from linearizing $\mathcal{F}$ at q^k and minimizing the Tikhonov functional of the linearization

$$q(\alpha_k) = \operatorname*{arg\,min}_{q \in \mathscr{Q}_{\text{ad}}} \frac{1}{2} \|\mathcal{F}(q^k) + \mathcal{F}'(q^k)(q - q^k) - y^\delta\|_{\mathcal{C}}^2 + \frac{\alpha_k}{2} \|q - q_\circ\|_{\mathscr{Q}}^2, \qquad (\mathbf{IP}_\alpha^k)$$

where $q_\circ \in \mathscr{Q}$ is the regularization center and $\alpha_k > 0$ a Tikhonov parameter.

Iteratively regularized Gauß-Newton method [Kaltenbacher, Neubauer, Scherzer 2008]

Let $q^k \in \mathscr{Q}_{\mathrm{ad}}$ be a given iterate. The IRGNM update scheme results from linearizing $\mathcal{F}$ at q^k and minimizing the Tikhonov functional of the linearization

$$q(\alpha_k) = \operatorname*{arg\,min}_{q \in \mathscr{Q}_{\mathrm{ad}}} \frac{1}{2} \|\mathcal{F}(q^k) + \mathcal{F}'(q^k)(q - q^k) - y^\delta\|_{\mathcal{C}}^2 + \frac{\alpha_k}{2} \|q - q_\circ\|_{\mathscr{Q}}^2, \qquad (\mathbf{IP}_\alpha^k)$$

where $q_\circ \in \mathscr{Q}$ is the regularization center and $\alpha_k > 0$ a Tikhonov parameter.

We accept the iterate $q^{k+1} := q(\alpha_k)$, if $\alpha_k > 0$ is chosen such that the following linearization condition holds for $0 < \theta < \Theta < 1$:

$$\theta J(q^k) \leq \|\mathcal{F}(q^k) + \mathcal{F}'(q^k)(q(\alpha_k) - q^k) - y^\delta\|_{\mathcal{C}}^2 \leq \Theta J(q^k).$$

Algorithm 1 (IRGNM)

Require: Noise level δ, discrepancy parameter $\tau > 0$, initial guess q^0, initial regularization $\alpha_0 > 0$, regularization center $q_\circ$.

```
Initialize k = 0.
while ‖F(q^k) − y^δ‖_H > τδ do
    Solve subproblem (IP_α^k) for q(α_k).
    while Linearization condition is not fulfilled do
        Increase (decrease) α_k, s.t. linearization condition is satisfied.
        Solve subproblem (IP_α^k) for q(α_k).
    end while
    Set q^{k+1} = q(α_k), k = k + 1.
end while
```

Parameter space reduction

Our goal is to approximate the infinite-dimensional parameter space by an adaptively constructed low-dimensional reduced space

$$\mathscr{Q}_r = \mathrm{span}\,\{\phi_1, \dots, \phi_{n_\mathscr{Q}}\} \subset \mathscr{Q}_{\mathrm{ad}} \qquad \text{with dimension } n_\mathscr{Q} \ll N_\mathscr{Q}.$$

For $q(t) = \sum_{i=1}^{n_\mathscr{Q}} \mathrm{q_i(t)}\phi_i \in \mathscr{Q}_r$ and a coercive operator A_r that is affine in q we obtain the representation

$$A_r(q(t)) = \bar{A}_r + \sum_{i=1}^{n_\mathscr{Q}} \mathrm{q_i(t)}\tilde{A}_r(\phi_i) \quad \text{for all } u, v \in V$$

This results in a low dimensional parametrization of the PDE, such that the adaptive RB-TR approach can be readily applied.

Parameter space reduction

Our goal is to approximate the infinite-dimensional parameter space by an adaptively constructed low-dimensional reduced space

$$\mathscr{Q}_r = \mathrm{span}\,\{\phi_1, \ldots, \phi_{n_{\mathscr{Q}}}\} \subset \mathscr{Q}_{\mathrm{ad}} \qquad \text{with dimension } n_{\mathscr{Q}} \ll N_{\mathscr{Q}}.$$

For $q(t) = \sum_{i=1}^{n_{\mathscr{Q}}} \mathsf{q}_{\mathsf{i}}(\mathsf{t})\phi_i \in \mathscr{Q}_r$ and a coercive operator A_r that is affine in q we obtain the representation

$$A_r(q(t)) = \bar{A}_r + \sum_{i=1}^{n_{\mathscr{Q}}} \mathsf{q}_{\mathsf{i}}(\mathsf{t})\tilde{A}_r(\phi_i) \quad \text{for all } u, v \in V$$

This results in a low dimensional parametrization of the PDE, such that the adaptive RB-TR approach can be readily applied.

Open question: How to build/enrich the reduced parameter space $\mathscr{Q}_r$?

IRGNM: Optimality System

Optimality System

Let $u = \mathcal{S}(q) \in V$, $q \in \mathscr{Q}_{\text{ad}}$. The update $\tilde{q} \in \mathscr{Q}_{\text{ad}}$ is optimal for $(\mathbf{IP}_\alpha^k)$ if and only if there exists linearized state $\tilde{u} \in \mathscr{V}$ and adjoint $\tilde{p} \in \mathscr{V}$ satisfying the optimality system f.a.a. $t \in (0,T)$, with initial and terminal conditions $\tilde{u}(0) = \tilde{p}(T) = 0$.

$$\langle \partial_t \tilde{u}(t) + A(q(t))\tilde{u}(t), v\rangle_{V',V} + \langle A(d(t))u(t), v\rangle_{V',V} = 0, \text{ for all } v \in V, \qquad \text{(primal state)}$$

$$\langle -\partial_t \tilde{p}(t) + A(q(t))'\tilde{p}(t), v\rangle_{V',V} + \langle \mathcal{C}(u(t) + \tilde{u}(t) - y^\delta(t)), \mathcal{C}v\rangle_C = 0, \text{ for all } v \in V, \qquad \text{(dual state)}$$

$$\underbrace{\int_0^T \langle A(e(t))u(t), \tilde{p}(t)\rangle_{V',V}\,\mathrm{d}t}_{=\langle \nabla \hat{J}(q), e\rangle_{\mathscr{Q}}} + \int_0^T \langle \alpha(\tilde{q}(t) - q_\circ(t)), e(t)\rangle_{\mathscr{Q}}\,\mathrm{d}t = 0, \text{ for all } q \in \mathscr{Q}. \qquad \text{(parameter)}$$

IRGNM: Optimality System

Optimality System

Let $u = \mathcal{S}(q) \in V$, $q \in \mathscr{Q}_{\text{ad}}$. The update $\tilde{q} \in \mathscr{Q}_{\text{ad}}$ is optimal for $(\mathbf{IP}_\alpha^k)$ if and only if there exists linearized state $\tilde{u} \in \mathscr{V}$ and adjoint $\tilde{p} \in \mathscr{V}$ satisfying the optimality system f.a.a. $t \in (0, T)$, with initial and terminal conditions $\tilde{u}(0) = \tilde{p}(T) = 0$.

$$
\begin{aligned}
\langle \partial_t \tilde{u}(t) + A(q(t))\tilde{u}(t), v\rangle_{V',V} + \langle A(d(t))u(t), v\rangle_{V',V} &= 0, \ \text{for all } v \in V, && \text{(primal state)}\\
\langle -\partial_t \tilde{p}(t) + A(q(t))'\tilde{p}(t), v\rangle_{V',V} + \langle \mathcal{C}(u(t) + \tilde{u}(t) - y^\delta(t)), \mathcal{C}v\rangle_C &= 0, \ \text{for all } v \in V, && \text{(dual state)}\\
\underbrace{\int_0^T \langle A(e(t))u(t), \tilde{p}(t)\rangle_{V',V}\,\mathrm{d}t}_{=\langle \nabla \hat{J}(q), e\rangle_{\mathscr{Q}}} + \int_0^T \langle \alpha(\tilde{q}(t) - q_\circ(t)), e(t)\rangle_{\mathscr{Q}}\,\mathrm{d}t &= 0, \ \text{for all } q \in \mathscr{Q}. && \text{(parameter)}
\end{aligned}
$$

$\Longrightarrow$ Adaptively enrich $\mathscr{Q}_r$ with the dominant POD modes of the gradients $\nabla \hat{J}(q_r^{(i)}))$!

Combined parameter and state space reduction

Let a reduced state space $\mathscr{V}_r$ of dimension $n_V \in \mathbb{N}$ and a reduced parameter $q \in \mathscr{Q}_r$ be given.
The combined state and parameter RB approximation of the state is then given as $u_r = S_r(q) \in \mathscr{V}_r$:

$$\frac{1}{\Delta t} M_{H,r}(u_r^k - u_r^{k-1}) + A_r(q^k)u_r^k = L_r^k \quad \forall k \in \mathbb{K},$$

The RB forward operator is defined as $\mathcal{F}_r := \mathcal{C} \circ \mathcal{S}_r$ and the RB discrepancy as

$$J_r(q) := \frac{1}{2} \|\mathcal{F}_r(q) - y^\delta\|_{\mathcal{C}}^2.$$

Hence, the reduced inverse problem reads

$$\min J_r(q) \quad \text{s.t.} \quad q \in \mathscr{Q}_r.$$

Reduced Basis Trust Region Methods for Inverse Parameter Identification

Theorem (A-posteriori error estimate for J)

Let $q_r \in Q^K_{\mathrm{ad},r}$ *and* $a_{q_r^k} \geq a > 0$ *be the q-dependent coercivity constant of* $A_h(q_r^k)$*. Then it holds*

$$|J_r(q_r) - J_h(q_r)| \leq \Delta^J(q_r),$$

where the right-hand side is given by

$$\Delta^J(q_r) := \left(\Delta t \sum_{k=1}^{K} \|r^k_{ad}(u_r, p_r; q_r)\|^2_{V_h'}\right)^{\frac{1}{2}} \frac{\Delta^{pr}(q_r)}{\sqrt{a_{q_r}}} + \frac{\|\mathcal{C}\|^2_{\mathcal{L}(V,C)}}{2a_{q_r}} \Delta^{pr}(q_r)^2 \text{ and}$$

$$\Delta^{pr}(q_r) := \left(\Delta t \sum_{k=1}^{K} \frac{1}{a_{q_r}} \|r^k_{pr}(u_r; q_r)\|^2_{V_h'}\right)^{\frac{1}{2}},$$

where $a_{q_r} := \min_{k \in \mathbb{K}} a_{q_r^k}$.

Parameter and state space error aware RB-TR-IRGNM

With reduced basis subproblem models $J_r^{(i)}$, $i = 0, 1, 2, \ldots$ based on iteratively enriched parameter and state spaces

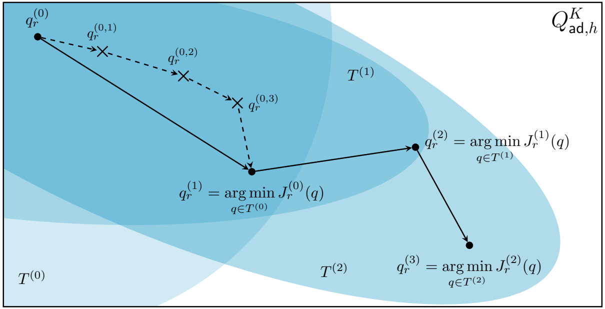
$$Q_r^{(0)} \subset Q_r^{(1)} \subset Q_r^{(2)} \subset \cdots \subset Q, \qquad V_r^{(0)} \subset V_r^{(1)} \subset V_r^{(2)} \subset \cdots \subset V.$$

TR subproblems in iteration i

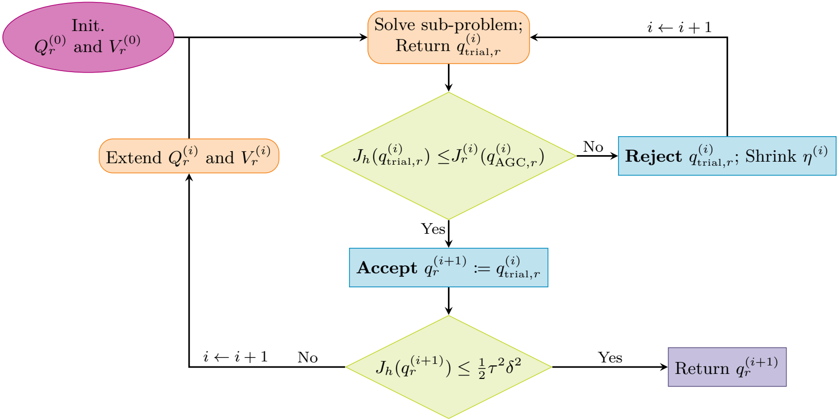
$$q_r = \underset{q \in \mathcal{Q}_r}{\arg\min}\, J_r^{(i)}(q) \quad \text{s.t.} \quad \frac{\Delta^{J^{(i)}}(q)}{|J_r^{(i)}(q)|} \leq \eta^{(i)}. \qquad (\mathbf{IP}_r^i)$$

Acceptance if sufficient decrease condition holds:

$$J_h(q_{\text{trial}\,r})^{(i)} \leq J_r^{(i)}(q_{\text{AGC},r}^{(i)}).$$



TR-IRGNM Algorithm



Example: Reconstruction of diffusion field

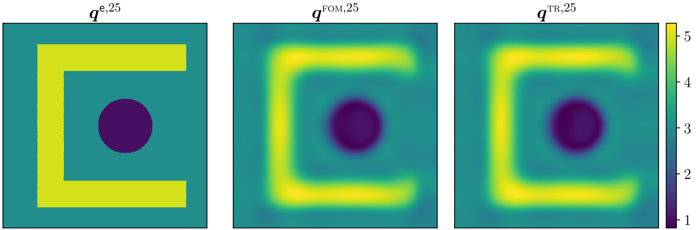
We consider the following parabolic PDE

$$\begin{aligned} \partial_t u(t,x) - \nabla(q(t,x)\cdot\nabla u(t,x)) &= f(t,x) \\ u(t,x) &= 0, \\ u(0,x) &= 0 \end{aligned} \quad .$$

which yields for $v, w \in V$:

$$\langle A(q(t))v, w\rangle_{V',V} := \int_\Omega q(t,x)\nabla v\cdot\nabla w \,\mathrm{d}x,$$

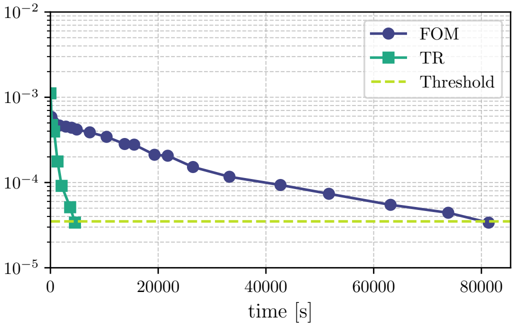
$$\langle l(t), v\rangle_{V',V} := \int_\Omega f(t)v \,\mathrm{d}x$$



Exact parameter at $k = 25$ (left) and FOM/RB-TR reconstructions.

Result of the Inverse Problem

Algo.	time [s]	FOM solves	n_Q	n_V	o. iter
FOM	81348	43695	–	–	21
TR	4548	14	68	197	6



Performance of the algorithms for the reconstruction of the diffusion coefficient (POD-Tol $1e-12$).

The overall speedup is **17,89**.

Application to a Structural Health Monitoring Model Problem

[Klein, Ohlberger, Schuster, 2026, arXiv:2605.19896]

Defect detection in a thin elastic plate

$$6.7\,\mathrm{mm} \times 1\,\mathrm{m} \times 1\,\mathrm{m}, \quad \Omega = [-0.1, 0.1] \times [-15, 15] \times [-15, 15], \quad \Gamma = \{y = \pm 15 \text{ or } z = \pm 15\}.$$

The displacement field $u : (0, T) \times \Omega \to \mathbb{R}^3$ solves

$$\begin{aligned} \rho\, \partial_t^2 u - \nabla \cdot \sigma_q(u) &= l && \text{in } (0, T) \times \Omega, \\ u &= 0 && \text{on } (0, T) \times \Gamma, \\ \sigma_q(u) \cdot n &= 0 && \text{on } (0, T) \times (\partial\Omega \setminus \overline{\Gamma}), \\ u(0, x) = 0, \partial_t u(0, x) &= 0 && \text{in } \Omega, \end{aligned}$$

with $\sigma_q(u) = \nabla_Y C_q(x, \nabla u)$ and

$$C_q(x, Y) = q(x) \left(\frac{\lambda}{2} [\mathrm{tr}(Y)]^2 + \mu \left\| \tfrac{1}{2}(Y + Y^\top) \right\|_F^2 \right), \quad q(x) = \begin{cases} \hat{q}(y, z), & x = -0.1, \\ 1, & \text{else.} \end{cases}$$

Application to a Structural Health Monitoring Model Problem

Excitation

$$\mathfrak{l}(t, x) = \mathfrak{l}_t(t)\, \mathfrak{l}_y(y)\, \mathfrak{l}_z(z)\, e_1$$

Admissible set

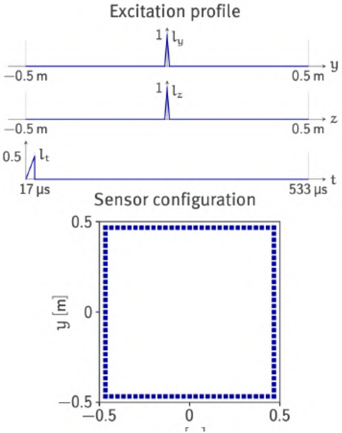
$$\mathcal{Q}_{ad} = \{\hat{q} \in L^2([-15, 15]^2) \mid \hat{q} > 0 \text{ a.e.}\}$$

FE-Discretization

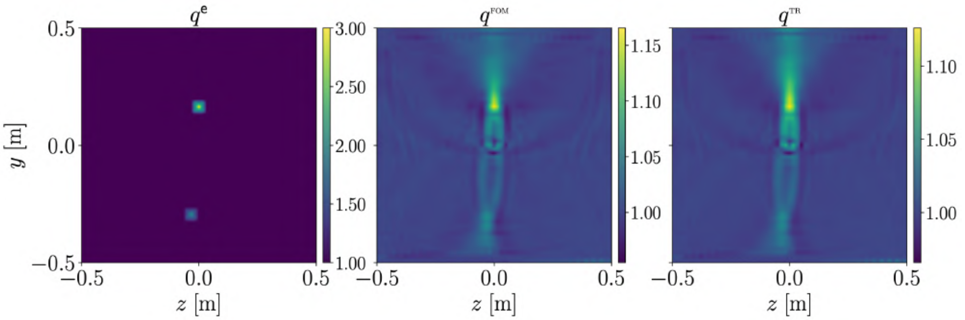
$$N_V = 55815, N_Q = 3721, K = 64$$

Noisy data

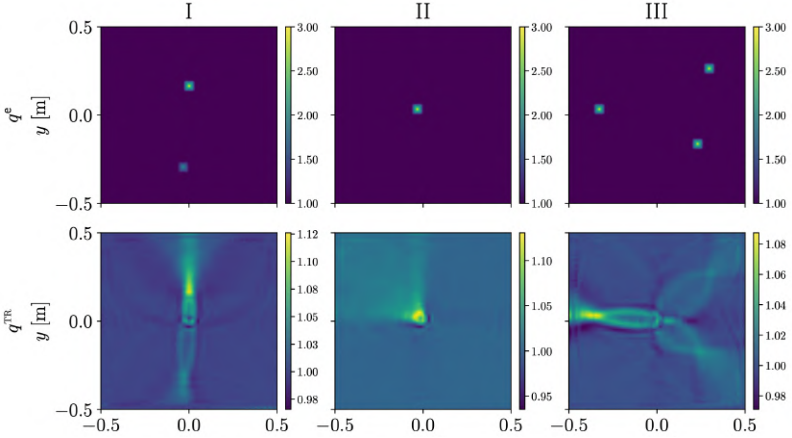
$$y^\delta = \mathcal{C}_h u_h^e + \delta \frac{\xi}{\|\xi\|_{C_h^K}}, \quad \delta = \delta_{\text{rel.}} \|\mathcal{C}_h u_h^e\|_{C_h^K}$$



Structural Health Monitoring: Point Defects; $\delta_{\mathrm{rel}} = 10^{-2}$



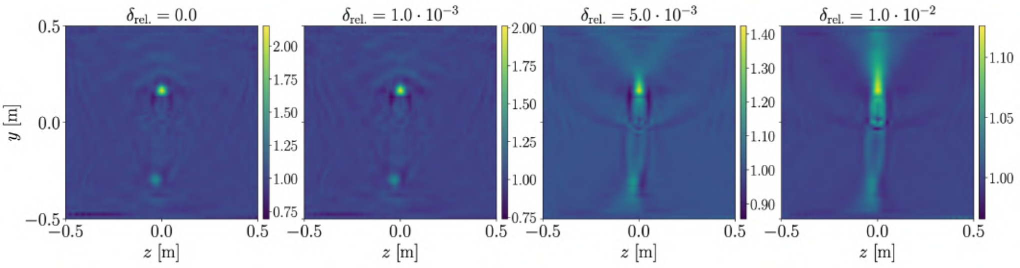
Structural Health Monitoring: Point Defects; $\delta_{\mathrm{rel}} = 10^{-2}$



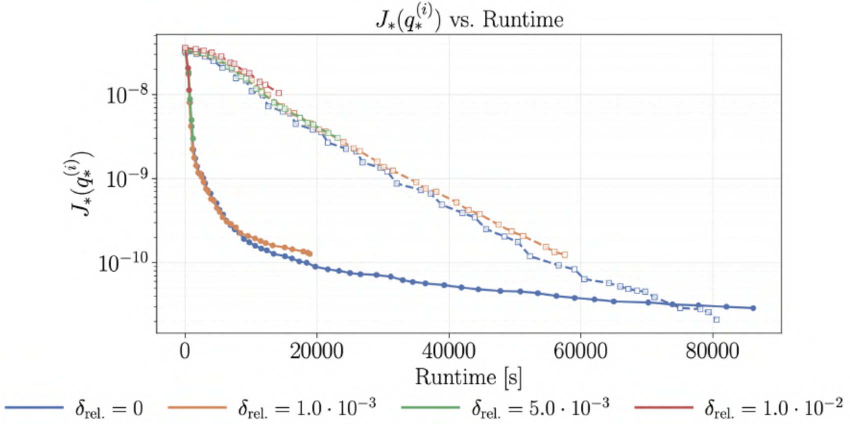
Structural Health Monitoring: Point Defects; $\delta_{\mathrm{rel}} = 10^{-2}$

Setup	Method	Q_h rel. err.	time [s]	speed-up	#FOM sol.
I	FOM	–	14238	–	466
	TR	2.75e-03	672	21.17	14
II	FOM	–	13425	–	472
	TR	3.38e-03	612	21.91	14
III	FOM	–	12674	–	420
	TR	2.02e-03	712	17.80	14

Structural Health Monitoring: Point Defects; varying noise level



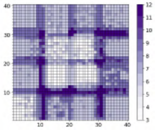
Structural Health Monitoring: Point Defects; varying noise level



Further recent contributions

- Parameter optimization for multiscale problems

$\Longrightarrow$ based on LOD multiscale method!



[Keil, Ohlberger. A Relaxed Localized Trust-Region Reduced Basis Approach for Optimization of Multiscale Problems. M2AN 2024]

$\Longrightarrow$ based on the LRBMS method!

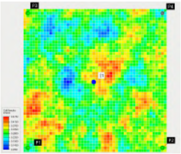
[Keil, Ohlberger, Schindler. Adaptive Localized Reduced Basis Methods for Large Scale PDE-constrained Optimization. LSSC 2024]

[Keil, Ohlberger, Schindler, Schleuß. Local training and enrichment based on a residual localization strategy, Algoritmy 2024]

Even more challenging problems

- Optimal control for large scale time dependent (non linear) PDE systems

  $\Longrightarrow$ adaptive machine learning based surrogate modeling



[Keil, Kleikamp, Lorentzen, Oguntola, Ohlberger. Adaptive machine learning based surrogate modeling to accelerate PDE-constrained optimization in enhanced oil recovery. ACOM 2022]

Thank you for your attention!

Thanks to my collaborators:

Bernard Haasdonk, Michael Kartmann, Hendrik Kleikamp,
Benedikt Klein, Stephan Rave, Felix Schindler,
Thomas Schuster, Stefan Volkwein, Tizian Wenzel

Software: DUNE, pyMOR

www.uni-muenster.de/MM/ohlberger