Universität Konstanz

DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Mixed−Integer Optimization
Optimization under Uncertainty
Optimization in Applications
Shape and Topology Optimization
Optimization and Machine Learning
2023 Heidelberg
EUCCO
6th European Conference on Computational Optimization
Optimal and Feedback Control of PDEs
Optimization on Manifolds
Non−smooth Optimization
Inverse Problems
Model−Order Reduction for Optimization

Adaptive Trust Region Reduced Basis Methods for Parameter Identification Problems

M. Kartmann, T. Keil, M. Ohlberger, S. Volkwein, B. Kaltenbacher

Universität Konstanz, September 27th, 2023

Outline

1 The parameter identification problem

1 The parameter identification problem

The parameter identification problem

Identify $q^{\mathrm{e}} \in \mathscr{Q}_{\mathrm{ad}} \subset \mathscr{Q}$ from

$$\mathcal{F}(q^{\mathrm{e}}) = y^{\mathrm{e}}, \qquad \textbf{(IP)}$$

where instead of the exact data y^{e} only noisy data $y^{\delta} \in \mathscr{H}$ is given with

$$\|y^{\mathrm{e}} - y^{\delta}\|_{\mathscr{H}} \leq \delta.$$

Forward operator $\mathcal{F} = \mathcal{C} \circ \mathcal{S}$

1. solution map $\mathcal{S} : \mathscr{Q}_{\mathrm{ad}} \to V$ with $\mathcal{S}(q) = u(q)$ and

$$a(u(q), v; q) = \ell(v) \quad \forall v \in V$$

2. observation map $\mathcal{C} : V \to \mathscr{H}$

$$y = \mathcal{C}u$$

- $\mathscr{Q}$ parameter space
- $\mathscr{Q}_{\mathrm{ad}} \subset \mathscr{Q}$ admissible parameters
- $\mathscr{H}$ observation space
- V state space

Goal: construct certified reduced forward operator $\mathcal{F}_r$.

1 The parameter identification problem - Assumptions

$$\mathcal{F}(q^{\mathrm{e}}) = y^{\mathrm{e}}, \quad \|y^{\mathrm{e}} - y^{\delta}\|_{\mathscr{H}} \leq \delta \tag{IP}$$

Assumptions

(i) Let $\mathscr{Q}, V, \mathscr{H}$ be real Hilbert spaces and $\mathscr{Q}_{\mathrm{ad}} \subset \mathscr{Q}$ be closed and convex.

(ii) For $q \in \mathscr{Q}_{\mathrm{ad}}$, $a(\cdot\,,\cdot\,;q)$ is a coercive and continuous, bilinear form and for $u, v \in V$ the map $q \mapsto a(u,v;q)$ is linear and $l \in V'$.

(iii) The solution map $\mathcal{S} : \mathscr{Q}_{\mathrm{ad}} \to V$ is non-linear and the observation map $\mathcal{C} : V \to \mathscr{H}$ is linear and bounded.

(iv) We assume $y^{\mathrm{e}} \in \mathrm{range}(\mathcal{F})$, $q^{\mathrm{e}} \in \mathring{\mathscr{Q}}_{\mathrm{ad}}$ and $\mathcal{F} : \mathscr{Q}_{\mathrm{ad}} \to \mathscr{H}$ to be injective, continuously Fréchet-differentiable and that (**IP**) is ill-posed in the sense that $\mathcal{F}$ is not continuously invertible.

- Simple inversion fails since $\mathcal{F}^{-1}(y^{\delta}) \not\to \mathcal{F}^{-1}(y^{\mathrm{e}}) = q^{\mathrm{e}}$ as $\delta \to 0$

$\Longrightarrow$ iterative regularization methods e.g. [Kaltenbacher, Neubauer, Scherzer 2008]

1 The parameter identification problem - Examples

$$\mathcal{F}(q^{\mathrm{e}}) = y^{\mathrm{e}}, \quad \|y^{\mathrm{e}} - y^{\delta}\|_{\mathscr{H}} \leq \delta \tag{IP}$$

Examples

Let Ω be a convex, bounded domain, $\mathscr{H} = L^2(\Omega)$, $V = H_0^1(\Omega)$, $\mathcal{C} = id_{V,\mathscr{H}}$, $f \in L^2(\Omega)$ and for $q_{\mathrm{a}} \in L^\infty(\Omega)$

$$\mathscr{Q}_{\mathrm{ad}} := \left\{ q \in \mathscr{Q} \,\middle|\, 0 < q_{\mathrm{a}} \leq q \text{ in } \Omega \text{ a.e.} \right\}.$$

1. Reconstruction of the reaction coefficient $\mathscr{Q} = L^2(\Omega)$:
$$-\Delta u + qu = f \quad \text{in } H^{-1}(\Omega),$$
2. Reconstruction of the diffusion coefficient $\mathscr{Q} = H^2(\Omega)$:
$$-\nabla \cdot \left(q \nabla u\right) = f \quad \text{in } H^{-1}(\Omega).$$

Goal: fast and stable reconstruction of q^{e} from (**IP**) by regularization methods.

Iteratively regularized Gauß-Newton method (IRGNM)

- the discrepancy: $\hat{J}(q) = \frac{1}{2}\|\mathcal{F}(q) - y^\delta\|^2_{\mathscr{H}}$.

Alg. 1: FOM IRGNM [Kaltenbacher, Kirchner, Veljovic 2014]

1. Iterative linearization of (**IP**) and Tikhonov regularization:

$$q(\alpha_k) = \arg\min_{q \in \mathscr{Q}} \frac{1}{2}\|\mathcal{F}(q^k) + \mathcal{F}'(q^k)(q - q^k) - y^\delta\|^2_{\mathscr{H}} + \frac{\alpha_k}{2}\|q - q_\circ\|^2_{\mathscr{Q}}, \qquad (\mathbf{IP}^k_\alpha)$$

2. Choice of regularization parameter: accept $q^{k+1} := q(\alpha_k)$ if

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

3. Early stopping according to the discrepancy principle: for $\tau > 1$

$$\|\mathcal{F}(q^k) - y^\delta\|_{\mathscr{H}} \leq \tau\delta \leq \|\mathcal{F}(q^{k-1}) - y^\delta\|_{\mathscr{H}}, \qquad k = 0, \dots, k_*(\delta, y^\delta).$$

- Local convergence: $k_*(\delta, y^\delta)$ is finite and $q^{k_*(\delta, y^\delta)} \to q^{\mathsf{e}}$ as $\delta \to 0$.

$\Rightarrow$ FOM IRGNM is a many-query context for the forward operator $\mathcal{F}$.

2 Adaptive reduced basis approximation

The reduced basis (RB) method - Challenges

- Discretization: $\mathscr{Q}_h = \text{span}\{\phi_1, \dots, \phi_{N_\mathscr{Q}}\}$, $q = \sum_{i=1}^{N_\mathscr{Q}} \mathrm{q}_i \phi_i \in \mathscr{Q}_h$ and

$$a(u, v; q) = \sum_{i=1}^{N_\mathscr{Q}} \mathrm{q}_i \, a(u, v; \phi_i) \quad \text{for all } u, v \in V.$$

1. Greedy RB state space reduction [Hesthaven et al. 2016,...]

- offline/online phase
- construct global RB V_r
- restricted to bounded, low-dim. $\mathscr{Q}_h$ ($N_\mathscr{Q} \leq 100$)

2. Adaptive RB state space reduction [Garmatter et al. 2016]

- construct local RB V_r along the optimization path
- online enrichment
- can handle larger $\mathscr{Q}_h$ ($N_\mathscr{Q} \approx 900$)

3. Adaptive parameter- and state space reduction

- Step 1: reduce affine components ($\mathscr{Q}_h \to \mathscr{Q}_r$).
- Step 2: reduce the state space ($V_h \to V_r$)

But here the parameter space dimension $N_\mathscr{Q}$ can be arbitrary large.

(i) No projection: An arbitrary number $N_\mathscr{Q}$ of affine components needs to be projected onto the reduced state space V_r.

(ii) No certification: The assembly of residual-based error estimates is infeasible because it involves the computation of about $\mathcal{O}(N_\mathscr{Q})$ Riesz representatives.

Step 1: Reduction of the parameter space

- $\mathscr{Q}_r = \text{span}\{\phi_1, \ldots, \phi_{n_\mathscr{Q}}\} \subset \mathscr{Q}$ with dimension $n_\mathscr{Q} \ll N_\mathscr{Q}$, $q = \sum_{i=1}^{n_\mathscr{Q}} \mathsf{q}_i \phi_i \in \mathscr{Q}_r$

$$a(u, v; q) = \sum_{i=1}^{n_\mathscr{Q}} \mathsf{q}_i a_i(u, v; \phi_i) \quad \text{for all } u, v \in V.$$

How to choose the snapshots ϕ_i?

Suppose $\bar{q} \in \mathscr{Q}_{\text{ad}}$ is an unconstrained local minimizer of the regularized discrepancy

$$\hat{J}_\alpha(q) = \hat{J}(q) + \frac{\alpha}{2} \|q - q_\circ\|^2_\mathscr{Q}, \qquad q, q_\circ \in \mathscr{Q}_{\text{ad}}, \alpha > 0,$$

then there exist $\bar{u}, \bar{p} \in V$ with

$$\begin{aligned} a(\bar{u}, v; \bar{q}) &= l(v) && \text{for all } v \in V, \\ a(v, \bar{p}; \bar{q}) &= -\langle \mathcal{C}\bar{u} - y^\delta, \mathcal{C}v \rangle_\mathscr{H} && \text{for all } v \in V, \\ \alpha(\bar{q} - q_\circ) + \mathcal{J}_\mathscr{Q}^{-1} \partial_q a(\bar{u}, \bar{p}; \cdot) &= 0, \end{aligned}$$

where $\mathcal{J}_\mathscr{Q} : \mathscr{Q} \to \mathscr{Q}'$ is the Riesz map. In particular, this implies

$$\bar{q} = q_\circ - \frac{1}{\alpha} \mathcal{J}_\mathscr{Q}^{-1} \partial_q a(\bar{u}, \bar{p}; \cdot).$$

$\Rightarrow$ Iterative parameter reduction $\mathscr{Q}_r^k = \text{span}\{q_\circ, q^0, \nabla\hat{J}(q^0), \ldots, \nabla\hat{J}(q^k)\}$.

$\Rightarrow$ Iterative state reduction $V_r^k = \text{span}\{u(q^0), p(q^0), \ldots, u(q^k), p(q^k)\}$.

Step 2: Reduction of the state space

- Given reduced spaces V_r, $\mathscr{Q}_r$ and $q = \sum_{i=1}^{n_{\mathscr{Q}}} \mathsf{q}_i \phi_i \in \mathscr{Q}_r$
- Reduced solution operator: $\mathcal{S}_r : \mathscr{Q}_r \cap \mathscr{Q}_{\mathrm{ad}} \to V_r$, $q \mapsto u_r$ with

$$\sum_{i=1}^{n_{\mathscr{Q}}} \mathsf{q}_i a_i(u_r, v; \phi_i) = \ell(v) \quad \text{for all } v \in V_r \tag{2.2}$$

- Reduced forward operator: $\mathcal{F}_r = \mathcal{C} \circ \mathcal{S}_r$
- Reduced discrepancy: $\hat{J}_r(q) := \frac{1}{2} \|\mathcal{F}_r(q) - y^\delta\|_{\mathscr{H}}^2$

Lemma (A-posteriori error estimate for $\hat{J}$)

Let $q \in \mathscr{Q}_r \cap \mathscr{Q}_{\mathrm{ad}}$ and $\underline{a}_q > 0$ be the coercivity constant for $a(\cdot, \cdot; q)$. Then:

$$|\hat{J}(q) - \hat{J}_r(q)| \leq \Delta_{\hat{J}}(q),$$

with

$$\Delta_{\hat{J}}(q) := \frac{\|\mathcal{C}\|_{\mathscr{L}(V,\mathscr{H})}^2}{2} \Delta_{pr}(q)^2 + \|r_{du}(u_r, p_r; q)\|_{V'} \Delta_{pr}(q),$$

$$\Delta_{pr}(q) := \frac{1}{\underline{a}_q} \|r_{pr}(u_r; q)\|_{V'},$$

where $r_{pr}(u_r; q)$, $r_{du}(u_r, p_r; q) \in V'$ are the primal and dual residual.

Trust Region Reduced Basis IRGNM ($\mathcal{Q}_r$-V_r-IRGNM)

Error-aware trust region framework [Qian et al. 17, Banholzer et al. 22]

1. Use $\hat{J}_r$ as a model function.
2. The trust region is defined by the relative error estimator

$$\mathcal{R}^k_{\hat{J}}(q) := \frac{\Delta_{\hat{J}}(q)}{\hat{J}_r(q)} \leq \eta^{(k)} \quad \text{for } q \in \mathcal{Q}_r^k \cap \mathcal{Q}_{\mathrm{ad}},$$

where $\eta^{(k)} > 0$ is the trust region radius at iteration k.

Alg. 2: Trust Region Reduced Basis IRGNM ($\mathcal{Q}_r$-V_r-IRGNM)

1. Initialize $V_r^0 = \mathrm{span}\{u(q^0), p(q^0)\}$, $\mathcal{Q}_r^0 = \mathrm{span}\{q_\circ, q^0, \nabla \hat{J}(q^0)\}$ by orthonormalization.
2. **while** $\|\mathcal{F}(q^k) - y^\delta\|_{\mathscr{H}} > \tau\delta$**:**
3. Compute AGC point $q^k_{\mathrm{AGC}} = q^k - t_k \nabla \hat{J}_r(q^k)$.
4. Solve reduced subproblem (**IP**$_r^k$) for q_{trial}

$$\min_{q \in \mathcal{Q}_r^k} \hat{J}_r(q) \quad \text{s.t.} \quad \mathcal{R}^k_{\hat{J}}(q) \leq \eta^{(k)}. \qquad (\mathbf{IP}_r^k)$$

5. Acceptance or rejection of q_{trial}?
6. **end while**
7. **Output:** reconstruction q^{k_*}, reduced spaces $\mathcal{Q}_r^{k_*}, V_r^{k_*}$.

$\mathscr{Q}_r$-V_r-IRGNM - Acceptance or rejection?

- sufficient for acceptance:

$$\hat{J}(q_{\text{trial}}) \leq \hat{J}_r(q_{\text{trial}}) + \Delta_{\hat{J}}(q_{\text{trial}}) < \hat{J}_r(q^k_{\text{AGC}}).$$

Accept $q^{k+1} = q_{\text{trial}}$, compute $u(q^{k+1})$, $p(q^{k+1})$ and $\nabla \hat{J}((q^{k+1})$ and enrich the spaces $\mathscr{Q}_r^{k+1}$, V_r^{k+1} by orthonormalization, modify radius $\eta^{(k+1)}$ and set $k = k + 1$.

- sufficient for rejection:

$$\hat{J}_r(q_{\text{trial}}) - \Delta_{\hat{J}}(q_{\text{trial}}) > \hat{J}_r(q^k_{\text{AGC}}).$$

Reject q_{trial} and diminish the radius $\eta^{(k)}$.

Alg. 2: Trust Region Reduced Basis IRGNM ($\mathscr{Q}_r$-V_r-IRGNM)

1. Initialize $V_r^0 = \text{span}\{u(q^0), p(q^0)\}$, $\mathscr{Q}_r^0 = \text{span}\{q_\circ, q^0, \nabla \hat{J}(q^0)\}$ by orthonormalization.
2. **while** $\|\mathcal{F}(q^k) - y^\delta\|_{\mathscr{H}} > \tau\delta$**:**
3. Compute AGC point $q^k_{\text{AGC}} = q^k - t_k \nabla \hat{J}_r(q^k)$.
4. Solve reduced subproblem (**IP**$_r^k$) for q_{trial}.
5. Acceptance or rejection of q_{trial}?
6. **end while**
7. **Output:** reconstruction q^{k_*}, reduced spaces $\mathscr{Q}_r^{k_*}, V_r^{k_*}$.

3 Numerical results

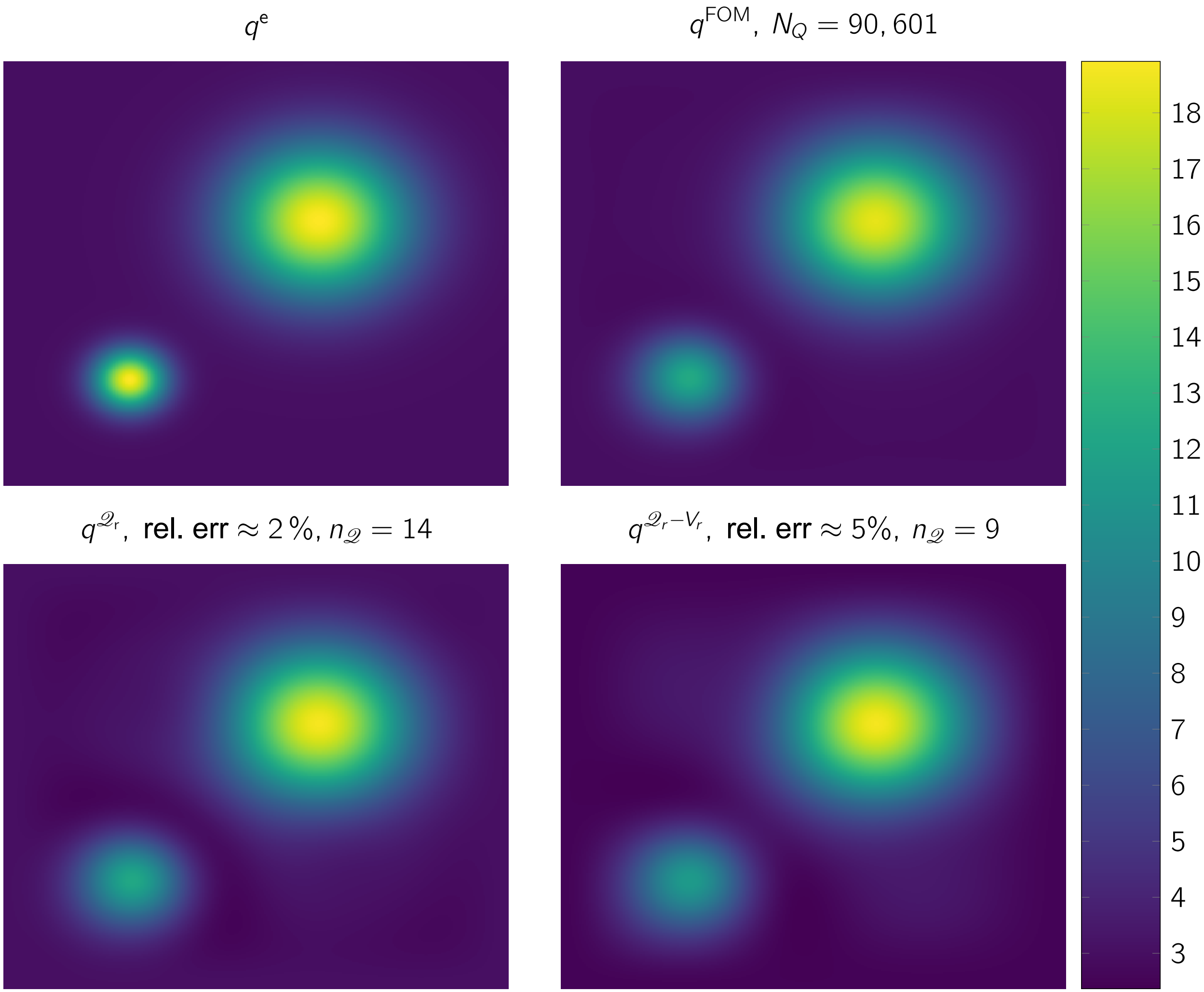
3 Numerical results

- Comparison of three algorithms:
  1. FOM-IRGNM: full-order model IRGNM (Alg. 1)
  2. $\mathscr{Q}_r$-V_r-IRGNM: parameter and state-reduced Trust Region IRGNM (Alg. 2)
  3. $\mathscr{Q}_r$-IRGNM: parameter-reduced IRGNM

- Problem setup:
  - Reaction ($\mathscr{Q} = L^2(\Omega)$) or diffusion case ($\mathscr{Q} = H^1(\Omega)$)
  - $\mathscr{H} = L^2(\Omega)$, $V = H_0^1(\Omega)$
  - $\Omega = (0, 1)^2$
  - FE discretization using $N_Q = 90,601$ dofs
  - background $q^0 = q_\circ \equiv 3 \in \mathscr{Q}_{\text{ad}}$
  - noise level $\delta = 10^{-5}$

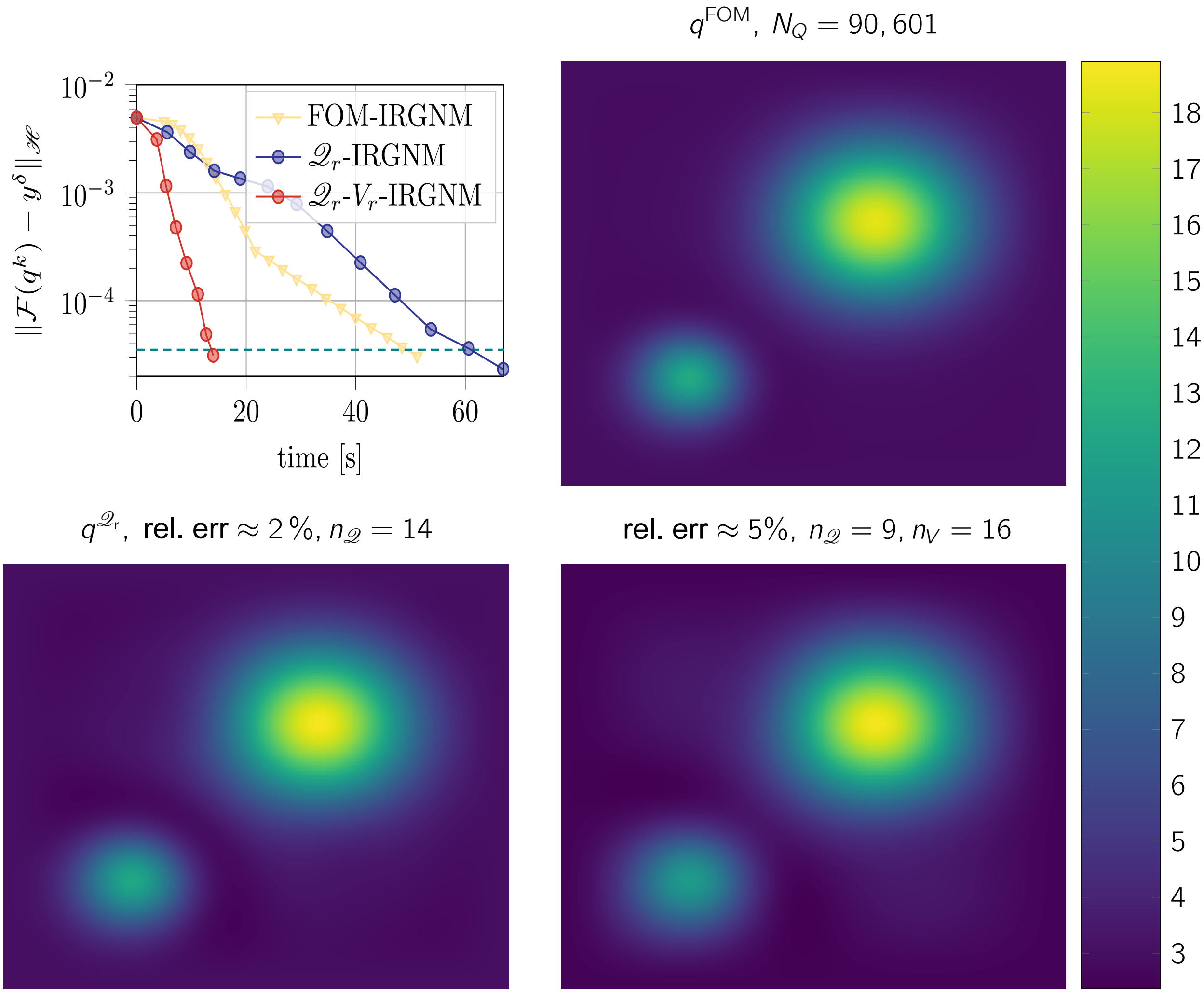
Reaction case

q^{e} from [Kirchner 2014]: $-\Delta u + qu = 1$ in $H^{-1}(\Omega)$



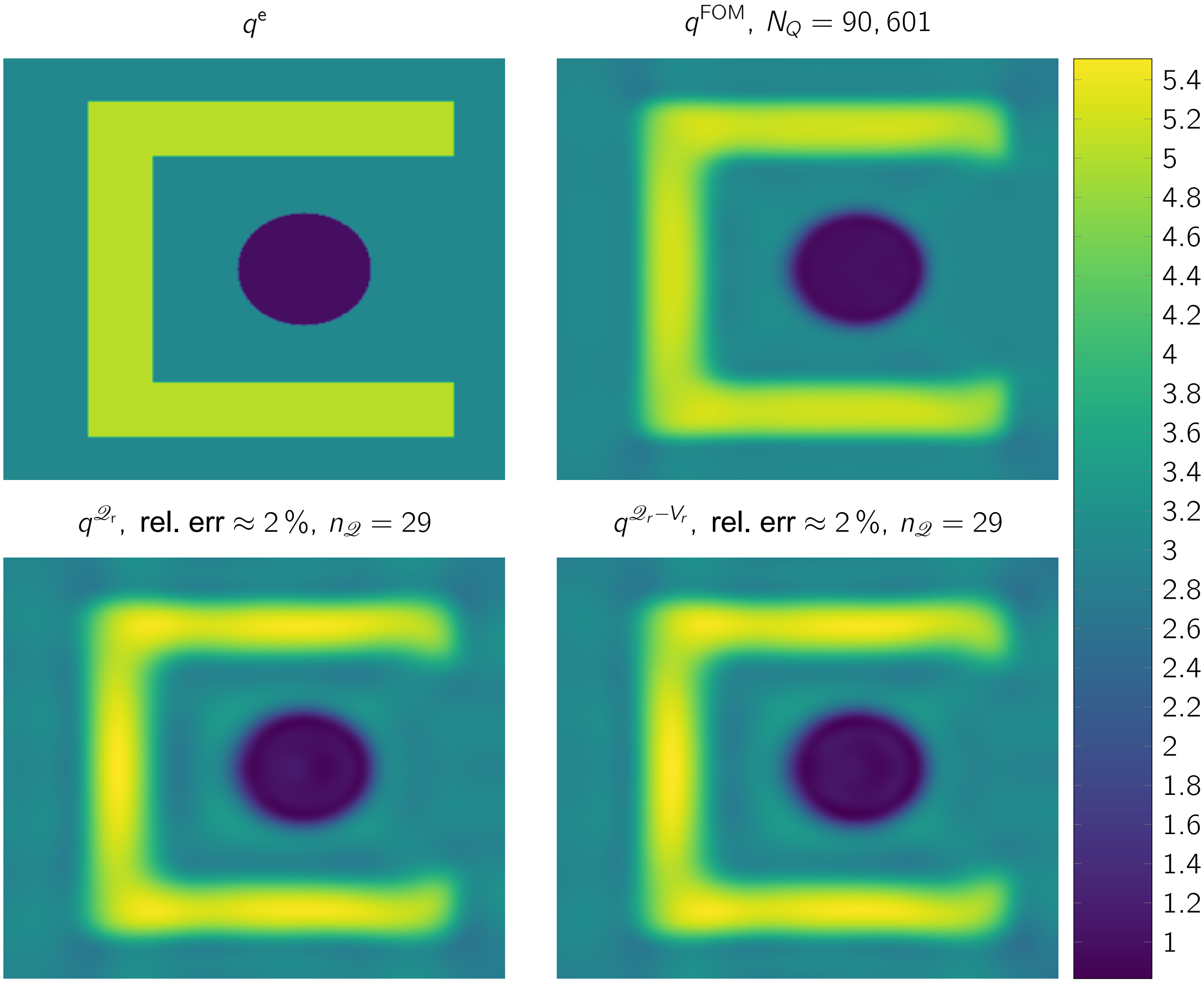
Reaction case

q^{e} from [Kirchner 2014]: $\quad -\Delta u + qu = 1$ in $H^{-1}(\Omega)$



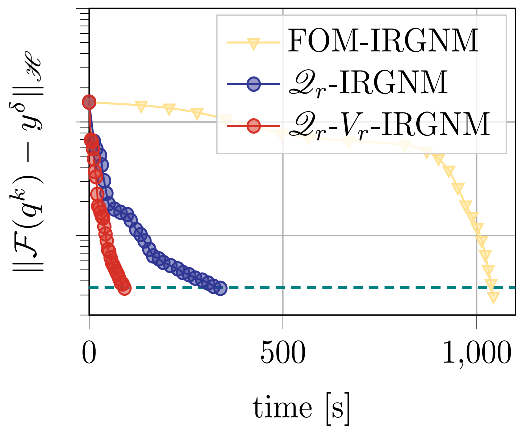
Diffusion case

q^{e} from [Garmatter et al. 16]: $-\nabla \cdot (q \nabla u) = 1$ in $H^{-1}(\Omega)$.

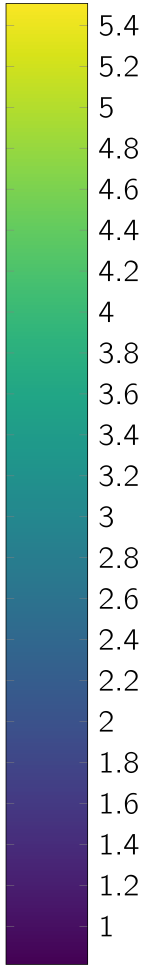


Diffusion case

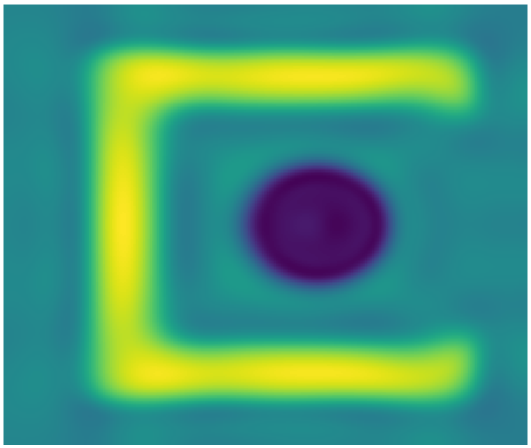
q^e from [Garmatter et al. 16]: $-\nabla \cdot (q \nabla u) = 1$ in $H^{-1}(\Omega)$.



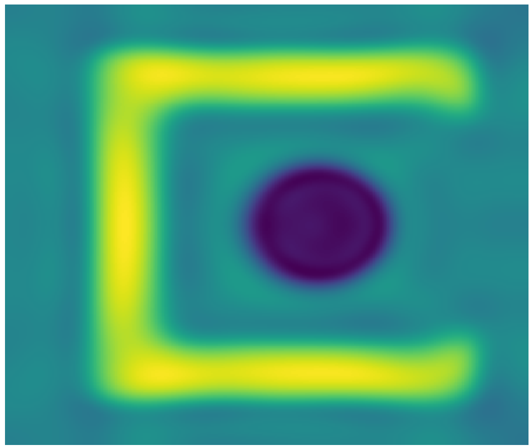
q^{FOM}, $N_Q = 90,601$



$q^{\mathscr{Q}_r}$, rel. err $\approx 2\,\%$, $n_{\mathscr{Q}} = 29$



rel. err $\approx 2\,\%$, $n_{\mathscr{Q}} = 29$, $n_V = 56$



4 Conclusion

- We introduced a new adaptive parameter and state space reduced regularization method based on RB model reduction.

- The enrichment of the reduced parameter space is motivated by the optimality condition of the regularized discrepancy, that transfers the low-rank structure of the (adjoint) state to the optimal parameter.

- The construction of the reduced parameter space allows us to apply the RB method to arbitrary dimension parameter spaces and to obtain fast low-dimensional reconstructions of the unknown parameter.

Bibliography

Banholzer, S. and Keil, T. and Mechelli, L. and Ohlberger, M. and Schindler, F. and Volkwein, S.
An adaptive projected Newton non-conforming dual approach for trust-region reduced basis approximation of PDE-constrained parameter optimization.
Pure and Applied Functional Analysis, 7:1561–1596, 2022.

Garmatter, D. and Haasdonk, B. and Harrach, B.
A reduced basis Landweber method for nonlinear inverse problems.
Inverse Problems, 32(3), 2016.

Kaltenbacher, B. and Kirchner, A. and Veljović, S.
Goal oriented adaptivity in the IRGNM for parameter identification in PDEs: I. reduced formulation.
Inverse Problems, 30(4), 2014.

Kartmann, M. and Keil, T. and Ohlberger, M. and Volkwein, S. and Kaltenbacher, B.
Adaptive Reduced Basis Trust Region Methods for Inverse Parameter Identification Problems.
Preprint (https://arxiv.org/pdf/2309.07627.pdf), 2023.