Universität Konstanz

Multiobjective Parameter Optimization of Elliptic PDEs using the Reduced Basis Method

S. Banholzer, B. Gebken, M. Dellnitz, S. Peitz, S. Volkwein

University of Konstanz / Paderborn University

Workshop: New trends in PDE constrained optimization

RICAM Linz, 17th October 2019

DFG Deutsche Forschungsgemeinschaft

SPP 1962

Problem Formulation

Consider the multiobjective parameter optimization problem

$$\min_{u,y} \begin{pmatrix} J_1(u,y) \\ \vdots \\ J_k(u,y) \end{pmatrix} \quad \text{s.t.} \quad e(y,u) = 0, \tag{MPOP}$$

where

- $J_i : U \times Y \to \mathbb{R}$, $i = 1, \ldots, k$, are the *objective functions*,
- $U = \mathbb{R}^n$ is the *parameter space*, Y is the *state space*,
- e is a PDE called the *state equation*, which is uniquely solvable for every $u \in U$. The parameter-to-state mapping is given by $\mathcal{S} : U \to Y$.
- We write $J_i(u) = J_i(u, \mathcal{S}(u))$.

Problem Formulation

Consider the multiobjective parameter optimization problem

$$\min_{u} \begin{pmatrix} J_1(u) \\ \vdots \\ J_k(u) \end{pmatrix} = \begin{pmatrix} J_1(u, \mathcal{S}(u)) \\ \vdots \\ J_k(u, \mathcal{S}(u)) \end{pmatrix}, \qquad \text{(MPOP)}$$

where

- $J_i : U \times Y \to \mathbb{R}$, $i = 1, \ldots, k$, are the *objective functions*,
- $U = \mathbb{R}^n$ is the *parameter space*, Y is the *state space*,
- e is a PDE called the *state equation*, which is uniquely solvable for every $u \in U$. The parameter-to-state mapping is given by $\mathcal{S} : U \to Y$.
- We write $J_i(u) = J_i(u, \mathcal{S}(u))$.

Problem Formulation

Consider the multiobjective parameter optimization problem

$$\min_u \begin{pmatrix} J_1(u) \\ \vdots \\ J_k(u) \end{pmatrix} = \begin{pmatrix} J_1(u, \mathcal{S}(u)) \\ \vdots \\ J_k(u, \mathcal{S}(u)) \end{pmatrix}, \qquad \text{(MPOP)}$$

where

- $J_i : U \times Y \to \mathbb{R}$, $i = 1, \dots, k$, are the *objective functions*,
- $U = \mathbb{R}^n$ is the *parameter space*, Y is the *state space*,
- e is a PDE called the *state equation*, which is uniquely solvable for every $u \in U$. The parameter-to-state mapping is given by $\mathcal{S} : U \to Y$.
- We write $J_i(u) = J_i(u, \mathcal{S}(u))$.

Questions:

- What is a meaningful optimality concept? (Problem: No total order on $\mathbb{R}^k$)

Problem Formulation

Consider the multiobjective parameter optimization problem

$$\min_{u} \begin{pmatrix} J_1(u) \\ \vdots \\ J_k(u) \end{pmatrix} = \begin{pmatrix} \mathcal{J}_1(u, \mathcal{S}(u)) \\ \vdots \\ \mathcal{J}_k(u, \mathcal{S}(u)) \end{pmatrix}, \qquad \text{(MPOP)}$$

where

- $\mathcal{J}_i : U \times Y \to \mathbb{R}$, $i = 1, \ldots, k$, are the *objective functions*,
- $U = \mathbb{R}^n$ is the *parameter space*, Y is the *state space*,
- e is a PDE called the *state equation*, which is uniquely solvable for every $u \in U$. The parameter-to-state mapping is given by $\mathcal{S} : U \to Y$.
- We write $J_i(u) = \mathcal{J}_i(u, \mathcal{S}(u))$.

Questions:

- What is a meaningful optimality concept? (Problem: No total order on $\mathbb{R}^k$)
- How can we solve (MPOP) (numerically)?

Problem Formulation

Consider the multiobjective parameter optimization problem

$$\min_u \begin{pmatrix} J_1^r(u) \\ \vdots \\ J_k^r(u) \end{pmatrix} = \begin{pmatrix} J_1(u, \mathcal{S}^r(u)) \\ \vdots \\ J_k(u, \mathcal{S}^r(u)) \end{pmatrix}, \tag{MPOP}$$

where

- $J_i : U \times Y \to \mathbb{R}$, $i = 1, \dots, k$, are the *objective functions*,
- $U = \mathbb{R}^n$ is the *parameter space*, Y is the *state space*,
- e is a PDE called the *state equation*, which is uniquely solvable for every $u \in U$. The parameter-to-state mapping is given by $\mathcal{S} : U \to Y$.
- We write $J_i(u) = J_i(u, \mathcal{S}(u))$.

Questions:

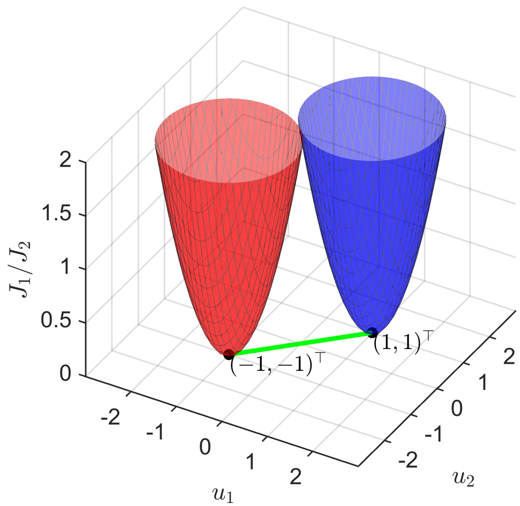
- What is a meaningful optimality concept? (Problem: No total order on $\mathbb{R}^k$)
- How can we solve (MPOP) (numerically)?
- How can approximation errors induced by using model order reduction be handled?

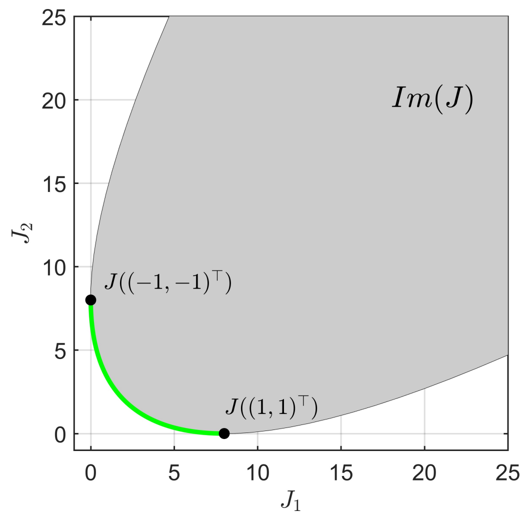
Pareto Optimality

Definition: Pareto optimality

A parameter $\bar{u} \in U$ is called *Pareto optimal* for (MPOP), if there is no other parameter $u \in U$ with

$$J_i(u) \leq J_i(\bar{u}) \quad \forall i \in \{1, \dots, k\},$$
$$J_l(u) < J_l(\bar{u}) \quad \text{for at least one } l \in \{1, \dots, k\}.$$





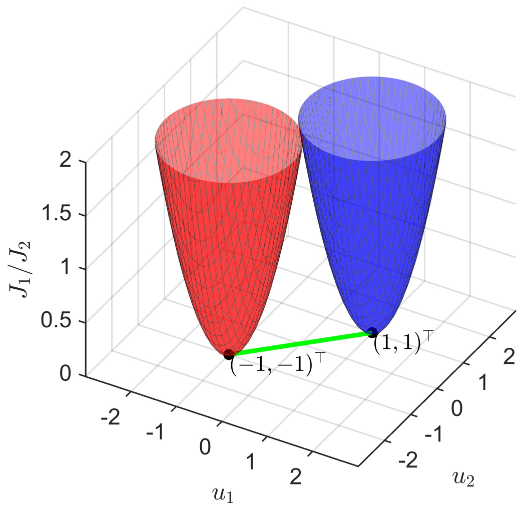
Example: $J : \mathbb{R}^2 \to \mathbb{R}^2, u \mapsto \begin{pmatrix} \|u - (-1,-1)^\top\|_2^2 \\ \|u - (1,1)^\top\|_2^2 \end{pmatrix}.$

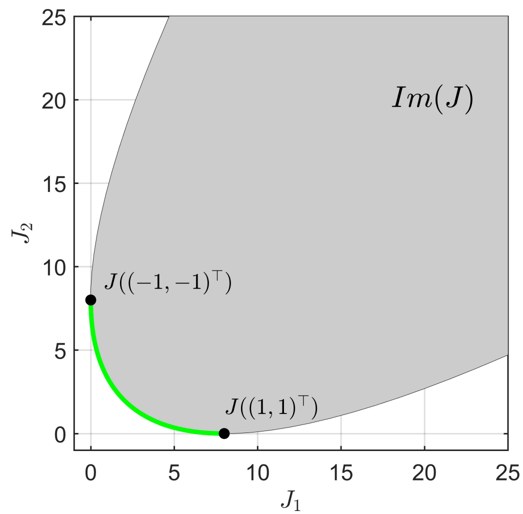
Pareto Optimality

Definition: Pareto optimality

A parameter $\bar{u} \in U$ is called *Pareto optimal* for (MPOP), if there is no other parameter $u \in U$ with

$$J_i(u) \leq J_i(\bar{u}) \quad \forall i \in \{1, \ldots, k\},$$
$$J_l(u) < J_l(\bar{u}) \quad \text{for at least one } l \in \{1, \ldots, k\}.$$





Goal: Find the *Pareto set* of (MPOP), i.e., the set of all Pareto optimal parameters of (MPOP).

Optimality conditions

Let J be differentiable.

Theorem:

If u is Pareto optimal, then there is some $\alpha \in (\mathbb{R}^{\geq 0})^k$ with $\sum_{i=1}^k \alpha_i = 1$ such that

$$\sum_{i=1}^{k} \alpha_i \nabla J_i(u) = DJ(u)^\top \alpha = 0. \tag{KKT}$$

Optimality conditions

Let J be differentiable.

Theorem:

If u is Pareto optimal, then there is some $\alpha \in (\mathbb{R}^{\geq 0})^k$ with $\sum_{i=1}^k \alpha_i = 1$ such that

$$\sum_{i=1}^{k} \alpha_i \nabla J_i(u) = DJ(u)^\top \alpha = 0. \qquad \text{(KKT)}$$

Definition: Pareto criticality

If u and α satisfy (KKT), then u is *Pareto critical* with *KKT vector* α. The set P_c of all Pareto critical points is the *Pareto critical set*.

Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.

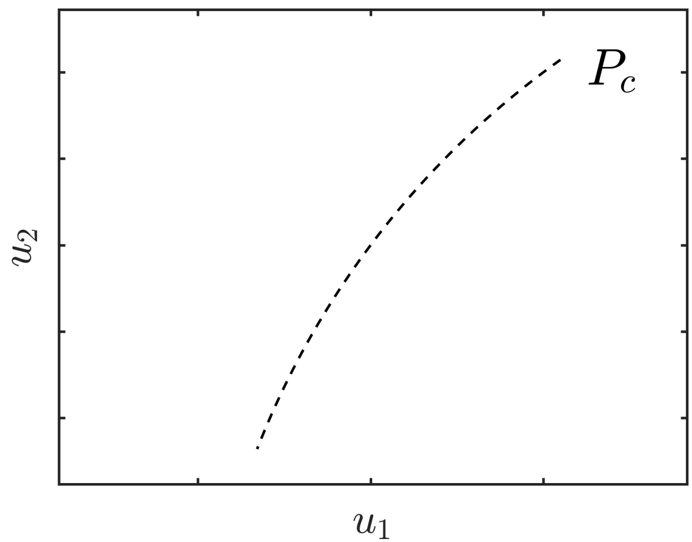
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



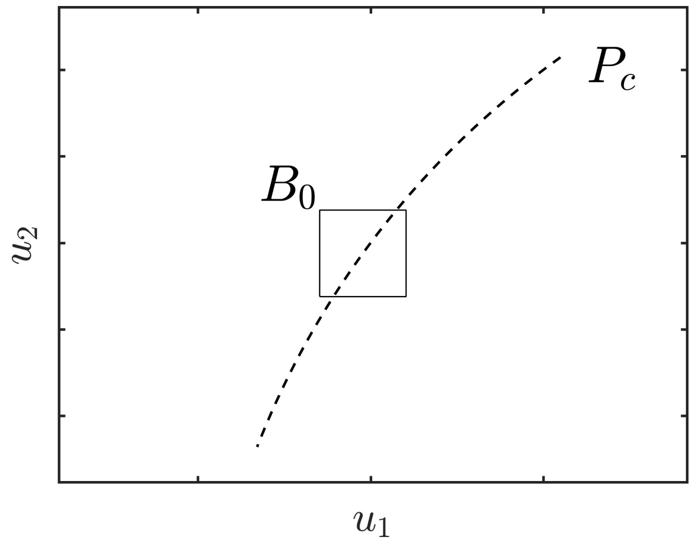
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



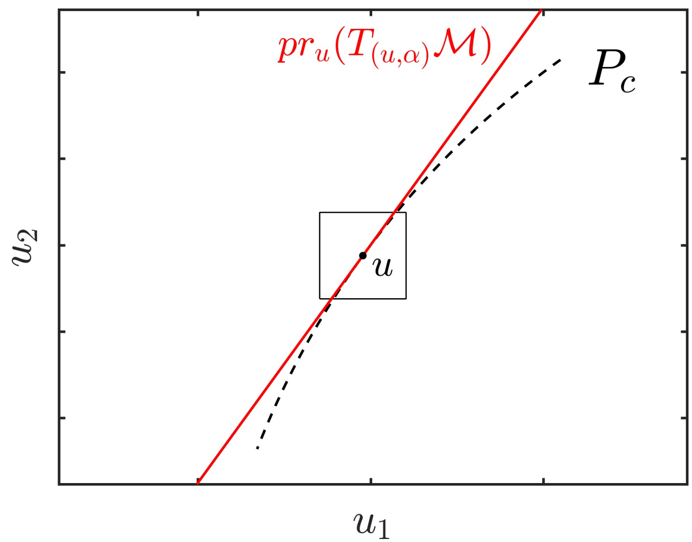
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \dots, 2i_n r)^T \mid (i_1, \dots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



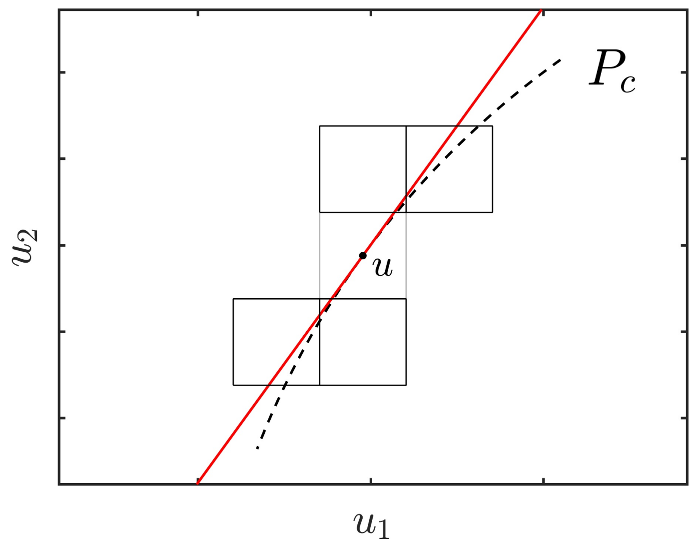
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \dots, 2i_n r)^T \mid (i_1, \dots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



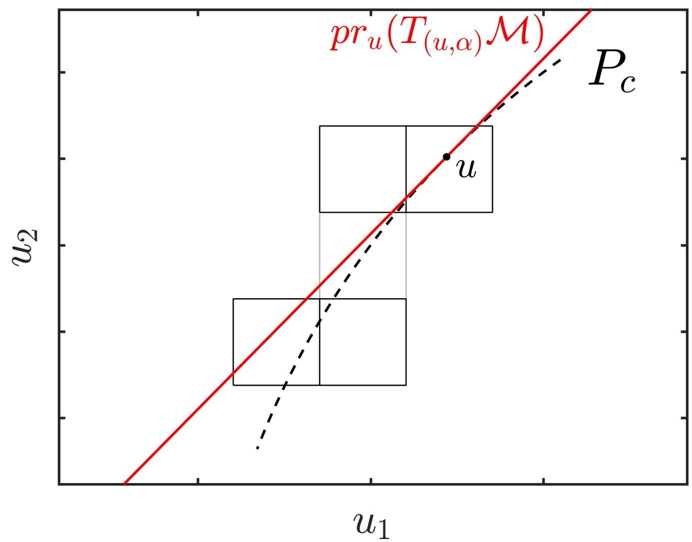
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



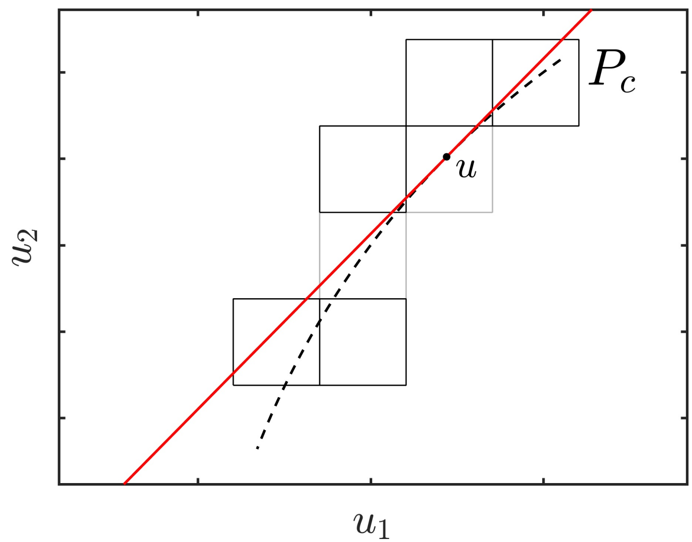
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



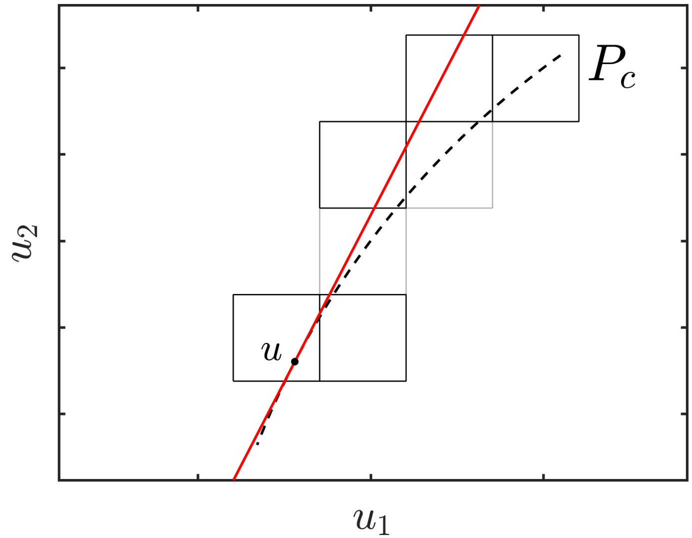
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\|DJ(u)^T \alpha\right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



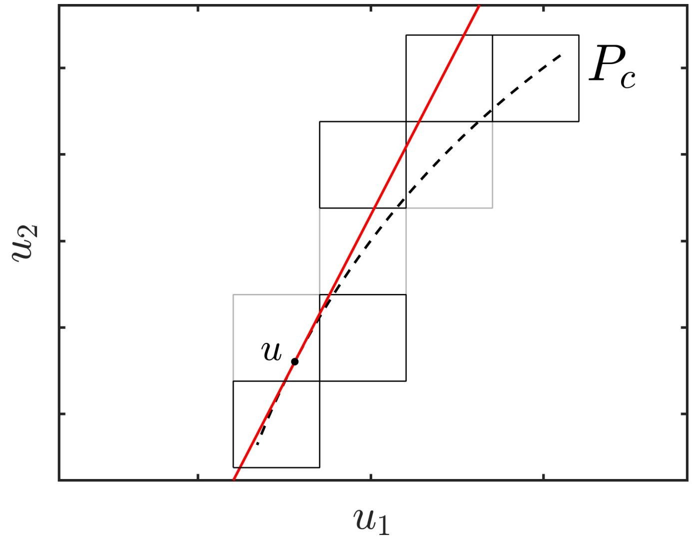
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



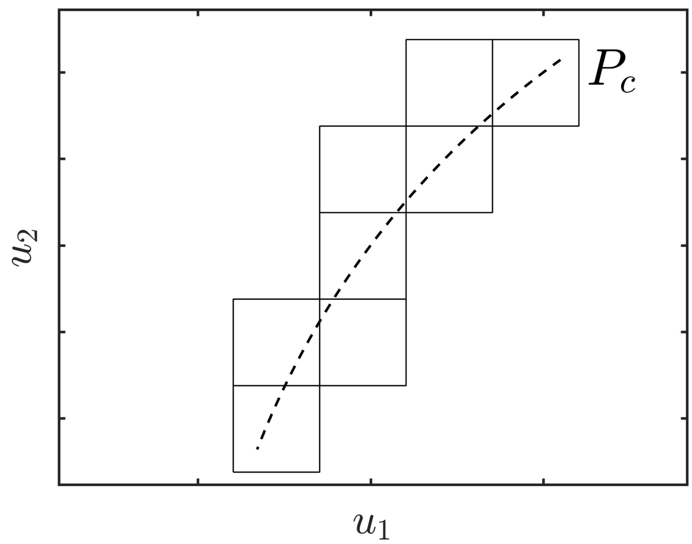
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \dots, 2i_n r)^T \mid (i_1, \dots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



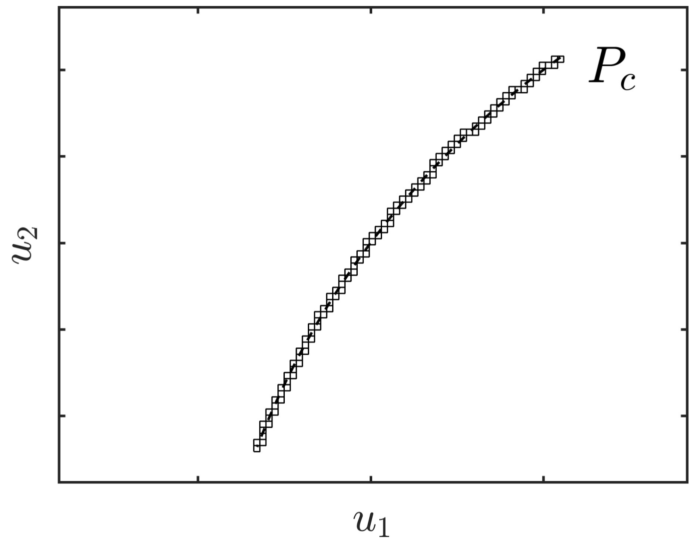
Continuation Method via Boxes

We use a set-oriented continuation method to compute the Pareto critical set. (see [Hillermeier, 2001])

- Divide the variable space into boxes with radius r

$$\mathcal{B}(r) := \{[-r, r]^n + (2i_1 r, \ldots, 2i_n r)^T \mid (i_1, \ldots, i_n) \in \mathbb{Z}^n\}.$$

- Compute $\mathcal{B}_c(r) := \{B \in \mathcal{B}(r) \mid B \cap P_c \neq \emptyset\}$.
- It holds $B \cap P_c \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2 = 0$.
- If a box $B \in \mathcal{B}(r)$ with $B \cap P_c \neq \emptyset$ is found, use the tangent space of P_c to get candidates for neighbouring boxes containing Pareto critical points.



Using inexact gradients

Assumption:

Let $J^r : U \to \mathbb{R}^k$ be the *inexact objective function* satisfying

$$\sup_{u \in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \;\; i \in \{1, \ldots, k\}$$

with *error bounds* $\varepsilon_1, \ldots, \varepsilon_k \geq 0$.

Using inexact gradients

Assumption:

Let $J^r : U \to \mathbb{R}^k$ be the *inexact objective function* satisfying

$$\sup_{u \in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \ \ i \in \{1, \ldots, k\}$$

with *error bounds* $\varepsilon_1, \ldots, \varepsilon_k \geq 0$.

Question: How can we approximatively compute P_c only using these information?

Using inexact gradients

Assumption:

Let $J^r : U \to \mathbb{R}^k$ be the *inexact objective function* satisfying

$$\sup_{u \in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \ \ i \in \{1, \ldots, k\}$$

with *error bounds* $\varepsilon_1, \ldots, \varepsilon_k \geq 0$.

Question: How can we approximatively compute P_c only using these information?

Lemma:

Let $u \in P_c$ with KKT vector $\alpha \in \Delta_k$. Then it holds

$$\|DJ^r(u)^\top \alpha\|_2 \leq \alpha^\top \varepsilon.$$

Using inexact gradients

Corollary:

Let

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Then $P_c^r \subseteq P^r$ and $P_c \subseteq P^r$. Furthermore, P^r is tight.

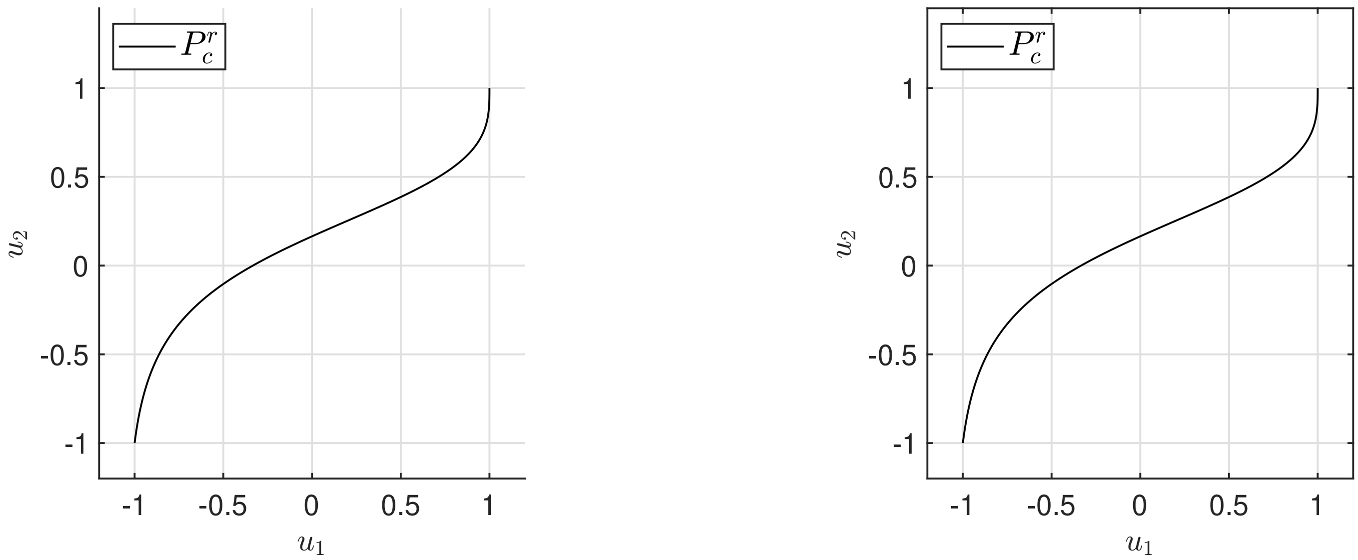
Using inexact gradients

Corollary:

Let

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Then $P_c^r \subseteq P^r$ and $P_c \subseteq P^r$. Furthermore, P^r is tight.



(a) $\varepsilon = (0.2, 0.05)^\top$ (b) $\varepsilon = (0, 0.2)^\top$

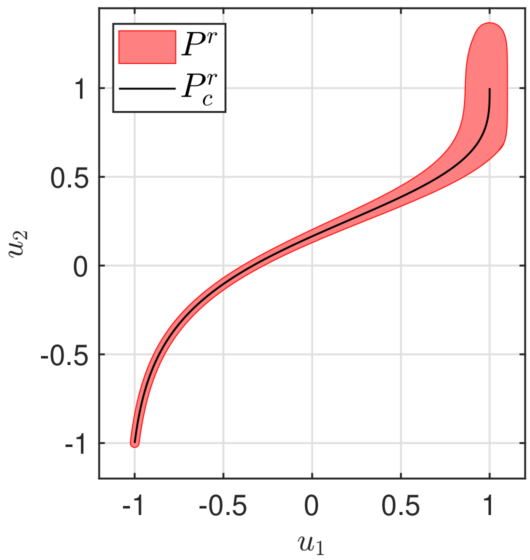
Using inexact gradients

Corollary:

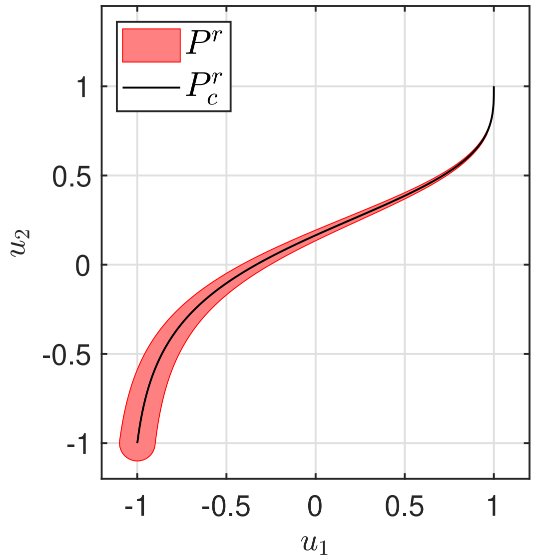
Let

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Then $P_c^r \subseteq P^r$ and $P_c \subseteq P^r$. Furthermore, P^r is tight.



(a) $\varepsilon = (0.2, 0.05)^\top$



(b) $\varepsilon = (0, 0.2)^\top$

Computing P^r

We want to compute a box covering of

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\},$$

i.e., $\{B \in \mathcal{B}(r) \mid B \cap P^r \neq \emptyset\}$.

Computing P^r

We want to compute a box covering of

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\},$$

i.e., $\{B \in \mathcal{B}(r) \mid B \cap P^r \neq \emptyset\}$.

Lemma/Problem:

P^r contains a nonempty, open subset of $\mathbb{R}^n$.

Computing P^r

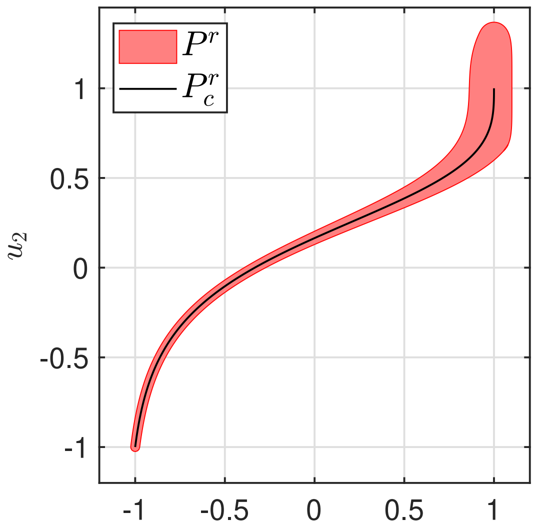
We want to compute a box covering of

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha\in\Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\},$$

i.e., $\{B \in \mathcal{B}(r) \mid B \cap P^r \neq \emptyset\}$.

Lemma/Problem:

P^r contains a nonempty, open subset of $\mathbb{R}^n$.



Two options:

- Strategy 1: Compute P^r directly.
- Strategy 2: Only compute ∂P^r.

Computing P^r

Let $B \in \mathcal{B}(r)$ and

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Computing P^r

Let $B \in \mathcal{B}(r)$ and

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Subproblem for Strategy 1 (Computing P^r):

$$B \cap P^r \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0$$

Computing P^r

Let $B \in \mathcal{B}(r)$ and

$$P^r := \{u \in \mathbb{R}^n \mid \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0\}.$$

Subproblem for Strategy 1 (Computing P^r):

$$B \cap P^r \neq \emptyset \iff \min_{u \in B, \alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2) \leq 0$$

Subproblem for Strategy 2 (Computing ∂P^r):

$$B \cap \partial P^r \neq \emptyset \iff \min_{u \in B} \varphi(u)^2 = 0$$

where

$$\varphi : \mathbb{R}^n \to \mathbb{R}, \quad u \mapsto \min_{\alpha \in \Delta_k} (\|DJ^r(u)^\top \alpha\|_2^2 - (\alpha^\top \varepsilon)^2).$$

Multiobjective Parameter Optimization of an Elliptic PDE

Given a domain $\Omega \subseteq \mathbb{R}^2$, we consider the problem

$$\min_{y,u} J(y,u) := \begin{pmatrix} \frac{1}{2}\|y - y^1\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|y - y^2\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|y - y^3\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|u\|^2_{\mathbb{R}^2} \end{pmatrix}$$

$$\begin{aligned} s.t. \quad -\kappa\Delta y(x) + cb(x)\cdot\nabla y(x) + 0.5y(x) &= f(x) && \text{for } x \in \Omega, \\ \frac{\partial y}{\partial \eta}(x) &= 0 && \text{for } x \in \partial\Omega, \\ u_a \leq u &\leq u_b. \end{aligned}$$

with $u = (\kappa, c) \in [0.5, 3] \times [-1, 1]$, i.e., we optimize the diffusivity κ in Ω and the strength and orientation c of the advection field b.

Multiobjective Parameter Optimization of an Elliptic PDE

Given a domain $\Omega \subseteq \mathbb{R}^2$, we consider the problem

$$\min_{y,u} J(y,u) := \begin{pmatrix} \frac{1}{2}\|y - y^1\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|y - y^2\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|y - y^3\|^2_{L^2(\Omega)} \\ \frac{1}{2}\|u\|^2_{\mathbb{R}^2} \end{pmatrix}$$

$$\begin{aligned} s.t. \quad -\kappa\Delta y(x) + cb(x)\cdot\nabla y(x) + 0.5y(x) &= f(x) && \text{for } x \in \Omega, \\ \frac{\partial y}{\partial \eta}(x) &= 0 && \text{for } x \in \partial\Omega, \\ u_a \le u &\le u_b. \end{aligned}$$

with $u = (\kappa, c) \in [0.5, 3] \times [-1, 1]$, i.e., we optimize the diffusivity κ in Ω and the strength and orientation c of the advection field b.

- Denote by $\mathcal{S}$ the solution operator of the state equation.
- We choose $y^1 = \mathcal{S}((0.7, 0.8))$, $y^2 = \mathcal{S}((2, 0.5))$, $y^3 = \mathcal{S}((3, -0.5))$.
- The solution operators of the adjoint equations (w.r.t. the cost functions J_1, J_2, J_3) are given by $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$.

Multiobjective Parameter Optimization of an Elliptic PDE

Given a domain $\Omega \subseteq \mathbb{R}^2$, we consider the problem

$$\min_u J(u) := \begin{pmatrix} \frac{1}{2}\|\mathcal{S}(u) - \mathcal{S}((0.7, 0.8))\|_{L^2(\Omega)}^2 \\ \frac{1}{2}\|\mathcal{S}(u) - \mathcal{S}((2, 0.5))\|_{L^2(\Omega)}^2 \\ \frac{1}{2}\|\mathcal{S}(u) - \mathcal{S}((3, -0.5))\|_{L^2(\Omega)}^2 \\ \frac{1}{2}\|u\|_{\mathbb{R}^2}^2 \end{pmatrix}$$
$$s.t. \quad u_a \leq u \leq u_b.$$

with $u = (\kappa, c) \in [0.5, 3] \times [-1, 1]$, i.e., we optimize the diffusivity κ in Ω and the strength and orientation c of the advection field b.

- Denote by $\mathcal{S}$ the solution operator of the state equation.
- We choose $y^1 = \mathcal{S}((0.7, 0.8))$, $y^2 = \mathcal{S}((2, 0.5))$, $y^3 = \mathcal{S}((3, -0.5))$.
- The solution operators of the adjoint equations (w.r.t. the cost functions J_1, J_2, J_3) are given by $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$.

Reduced Basis Method

For solving (MPOP) by the continuation method the problem

$$\min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2$$

has to be solved for numerous boxes B.

Reduced Basis Method

For solving (MPOP) by the continuation method the problem

$$\min_{u \in B, \alpha \in \Delta_k} \left\| DJ(u)^T \alpha \right\|_2^2$$

has to be solved for numerous boxes B.

→ Many solves of the state and adjoint equations are required.

→ Use the Reduced Basis method to lower the computational effort.

Reduced Basis Method

Weak formulation of the state equation: Find $y \in V := H^1(\Omega)$ such that

$$a(u; y, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V.$$

Reduced Basis Method

Weak formulation of the state equation: Find $y \in V := H^1(\Omega)$ such that

$$a(u; y, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V.$$

Idea of the RB method:

Replace the infinite-dimensional space V in the weak formulation by a low-dimensional subspace $V^r \subset V$, which is typically spanned by snapshots, i.e., solutions to the state (and adjoint) equation to different parameter values.

Reduced Basis Method

FE discretization of the state equation: Find $y^N \in V^N \subset H^1(\Omega)$ such that

$$a(u; y, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V^N.$$

Idea of the RB method:

Replace the FE space V^N in the FE discretization by a low-dimensional subspace $V^r \subset V^N$, which is typically spanned by snapshots, i.e., solutions to the state (and adjoint) equation to different parameter values.

Reduced Basis Method

FE discretization of the state equation: Find $y^N \in V^N \subset H^1(\Omega)$ such that

$$a(u; y, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V^N.$$

Idea of the RB method:

Replace the FE space V^N in the FE discretization by a low-dimensional subspace $V^r \subset V^N$, which is typically spanned by snapshots, i.e., solutions to the state (and adjoint) equation to different parameter values.

$\rightarrow$ RB state equation: Find $y^r \in V^r$ such that

$$a(u; y^r, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V^r.$$

Reduced Basis Method

FE discretization of the state equation: Find $y^N \in V^N \subset H^1(\Omega)$ such that

$$a(u; y, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V^N.$$

Idea of the RB method:

Replace the FE space V^N in the FE discretization by a low-dimensional subspace $V^r \subset V^N$, which is typically spanned by snapshots, i.e., solutions to the state (and adjoint) equation to different parameter values.

$\rightarrow$ RB state equation: Find $y^r \in V^r$ such that

$$a(u; y^r, \varphi) = b(\varphi) \quad \text{for all } \varphi \in V^r.$$

- Denote the solution operator of the RB state equation by $\mathcal{S}^r$.
- Let $\mathcal{A}_1^r, \mathcal{A}_2^r, \mathcal{A}_3^r$ be the solution operators of the RB adjoint equations.
- Define the RB objective function $J^r(u) := J(\mathcal{S}^r(u))$.

Constructing the Reduced Basis

To apply the strategies for the continuation method with inexact gradients, the RB has to be constructed such that the inequalities

$$\sup_{u \in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \quad \text{for all } i \in \{1, \ldots, k\}$$

hold for given error thresholds $\varepsilon_1, \ldots, \varepsilon_4 \geq 0$.
$\rightarrow$ Use a greedy algorithm.

Constructing the Reduced Basis

To apply the strategies for the continuation method with inexact gradients, the RB has to be constructed such that the inequalities

$$\sup_{u \in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \quad \text{for all } i \in \{1, \dots, k\}$$

hold for given error thresholds $\varepsilon_1, \dots, \varepsilon_4 \geq 0$.
$\rightarrow$ Use a greedy algorithm.

Data: Parameter set $\mathcal{P} \subset [0.5, 3] \times [-1, 1]$, greedy tolerances $\varepsilon_1, \dots, \varepsilon_k > 0$.

```
Choose u ∈ P, compute S(u), A_1(u), A_2(u), A_3(u);
Set V^r = span{S(u), A_1(u), A_2(u), A_3(u)} and compute the reduced
 basis by orthonormalization;
while max_{u∈P} max_{i∈{1,2,3,4}} ||∇J_i(u) − ∇J_i^r(u)||_2 > ε_i do
    Choose (ū, i) = arg max_{u∈P, i∈{1,2,3,4}} ||∇J_i(u) − ∇J_i^r(u)||_2;
    Compute S(ū) and A_i(ū);
    Set V^r = span{V^r ∪ {S(ū), A_i(ū)}} and compute the reduced basis by
     orthonormalization
```

Constructing the Reduced Basis

To apply the strategies for the continuation method with inexact gradients, the RB has to be constructed such that the inequalities

$$\sup_{u\in U} \|\nabla J_i(u) - \nabla J_i^r(u)\|_2 \leq \varepsilon_i, \quad \text{for all } i \in \{1, \ldots, k\}$$

hold for given error thresholds $\varepsilon_1, \ldots, \varepsilon_4 \geq 0$.

$\rightarrow$ Use a greedy algorithm.

Data: Parameter set $\mathcal{P} \subset [0.5, 3] \times [-1, 1]$, greedy tolerances $\varepsilon_1, \ldots, \varepsilon_k > 0$.

```
Choose u ∈ P, compute S(u), A_1(u), A_2(u), A_3(u);
Set V^r = span{S(u), A_1(u), A_2(u), A_3(u)} and compute the reduced
 basis by orthonormalization;
while max_{u∈P} max_{i∈{1,2,3,4}} ||∇J_i(u) − ∇J_i^r(u)||_2 > ε_i do
    Choose (ū, i) = arg max_{u∈P, i∈{1,2,3,4}} ||∇J_i(u) − ∇J_i^r(u)||_2;
    Compute S(ū) and A_i(ū);
    Set V^r = span{V^r ∪ {S(ū), A_i(ū)}} and compute the reduced basis by
     orthonormalization
```

Choose $u \in \mathcal{P}$, compute $\mathcal{S}(u), \mathcal{A}_1(u), \mathcal{A}_2(u), \mathcal{A}_3(u)$;

Set $V^r = span\{\mathcal{S}(u), \mathcal{A}_1(u), \mathcal{A}_2(u), \mathcal{A}_3(u)\}$ and compute the reduced basis by orthonormalization;

while $\max_{u\in\mathcal{P}} \max_{i\in\{1,2,3,4\}} \left\|\nabla J_i(u) - \nabla J_i^r(u)\right\|_2 > \varepsilon_i$ **do**

- Choose $(\bar{u}, i) = \arg\max_{u\in\mathcal{P},\, i\in\{1,2,3,4\}} \left\|\nabla J_i(u) - \nabla J_i^r(u)\right\|_2$;
- Compute $\mathcal{S}(\bar{u})$ and $\mathcal{A}_i(\bar{u})$;
- Set $V^r = span\{V^r \cup \{\mathcal{S}(\bar{u}), \mathcal{A}_i(\bar{u})\}\}$ and compute the reduced basis by orthonormalization

A-posteriori error estimation

It can be shown

$$|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u) \rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u) \rangle|$$

A-posteriori error estimation

It can be shown

$$
\begin{aligned}
&|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u)\rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u)\rangle| \\
\leq &\|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\|
\end{aligned} \tag{1}
$$

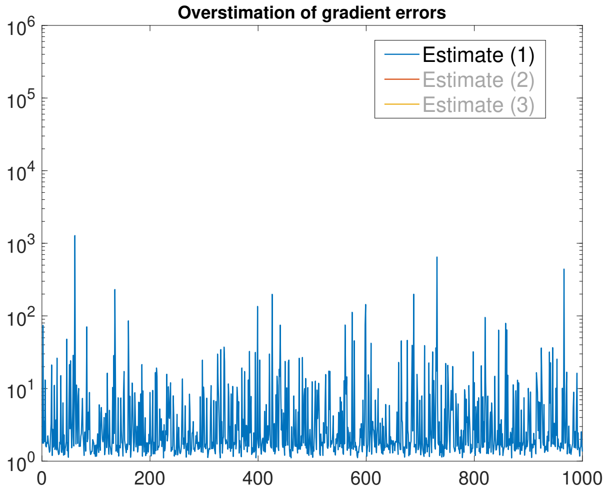


Figure: Overestimation of gradient errors

A-posteriori error estimation

It can be shown

$$
\begin{aligned}
&|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u)\rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u)\rangle| \\
\leq &\|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (1) \\
\leq &\|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u)\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (2)
\end{aligned}
$$

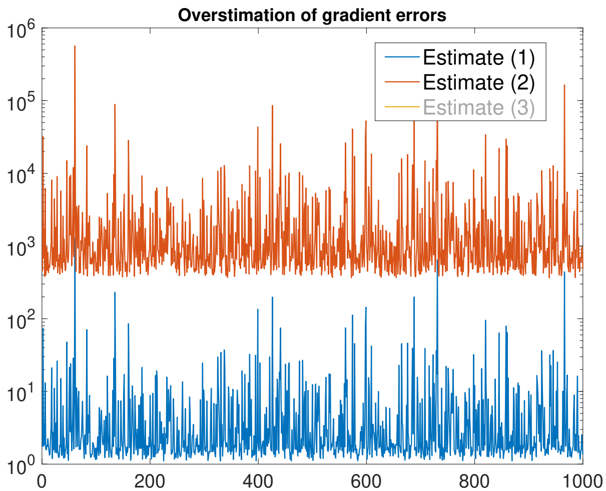


Figure: Overestimation of gradient errors

A-posteriori error estimation

It can be shown

$$
\begin{aligned}
&|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u)\rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u)\rangle| \\
\leq& \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (1) \\
\leq& \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u)\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (2) \\
\leq& \Delta_{\mathcal{S}}(u) \, \|\nabla \mathcal{A}_i^r(u)\| + \Delta_{\mathcal{S}}(u)\Delta_{\mathcal{A}_i}(u) + \|\nabla \mathcal{S}^r(u)\| \, \Delta_{\mathcal{A}_i}(u). \qquad (3)
\end{aligned}
$$

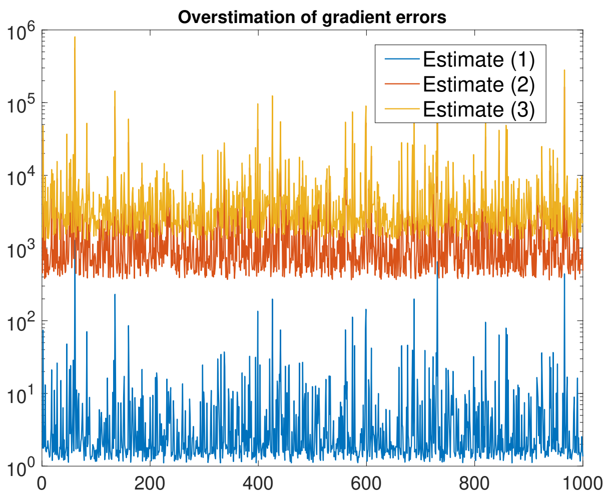


Figure: Overestimation of gradient errors

A-posteriori error estimation

It can be shown

$$
\begin{aligned}
&|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u)\rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u)\rangle| \\
\leq & \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
& + \|\nabla \mathcal{S}^r(u) \cdot \nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (1) \\
\leq & \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla \mathcal{A}_i^r(u)\| + \|\nabla (\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
& + \|\nabla \mathcal{S}^r(u)\| \, \|\nabla (\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (2) \\
\leq & \Delta_{\mathcal{S}}(u) \, \|\nabla \mathcal{A}_i^r(u)\| + \Delta_{\mathcal{S}}(u) \Delta_{\mathcal{A}_i}(u) + \|\nabla \mathcal{S}^r(u)\| \, \Delta_{\mathcal{A}_i}(u). \qquad (3)
\end{aligned}
$$

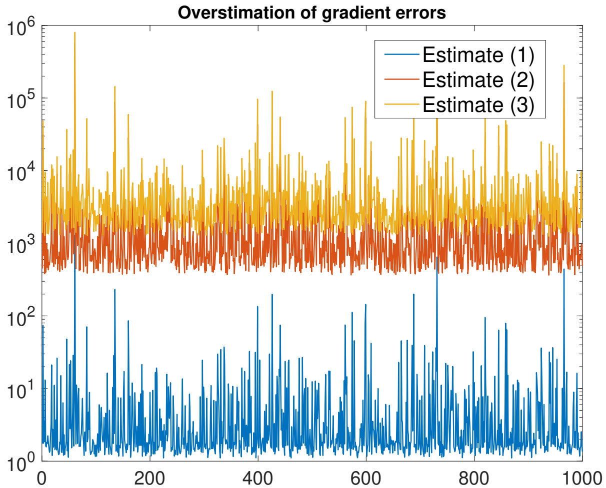


Figure: Overestimation of gradient errors

→ Huge overestimation of the gradient error (mainly caused by using the Cauchy-Schwarz inequality in (2)).

A-posteriori error estimation

It can be shown

$$
\begin{aligned}
&|\partial_\kappa J_i(u) - \partial_\kappa J_i^r(u)| = |\langle \nabla \mathcal{S}(u), \nabla \mathcal{A}_i(u)\rangle - \langle \nabla \mathcal{S}^r(u), \nabla \mathcal{A}_i^r(u)\rangle| \\
\leq &\|\nabla(\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla \mathcal{A}_i^r(u)\| + \|\nabla(\mathcal{S}(u) - \mathcal{S}^r(u)) \cdot \nabla(\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u) \cdot \nabla(\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (1) \\
\leq &\|\nabla(\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla \mathcal{A}_i^r(u)\| + \|\nabla(\mathcal{S}(u) - \mathcal{S}^r(u))\| \, \|\nabla(\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \\
&+ \|\nabla \mathcal{S}^r(u)\| \, \|\nabla(\mathcal{A}_i(u) - \mathcal{A}_i^r(u))\| \qquad (2) \\
\leq &\Delta_{\mathcal{S}}(u) \, \|\nabla \mathcal{A}_i^r(u)\| + \Delta_{\mathcal{S}}(u)\Delta_{\mathcal{A}_i}(u) + \|\nabla \mathcal{S}^r(u)\| \, \Delta_{\mathcal{A}_i}(u). \qquad (3)
\end{aligned}
$$

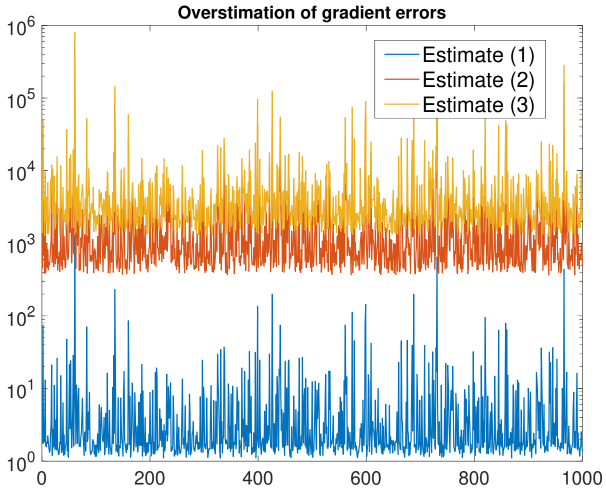


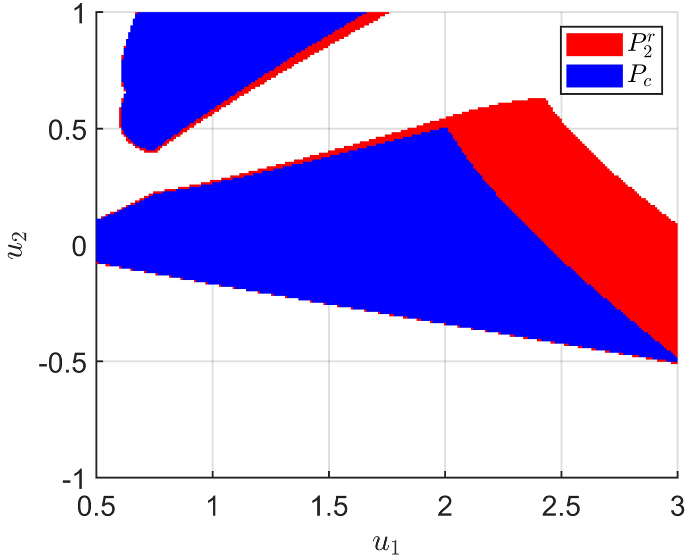
Figure: Overestimation of gradient errors

→ Huge overestimation of the gradient error (mainly caused by using the Cauchy-Schwarz inequality in (2)).

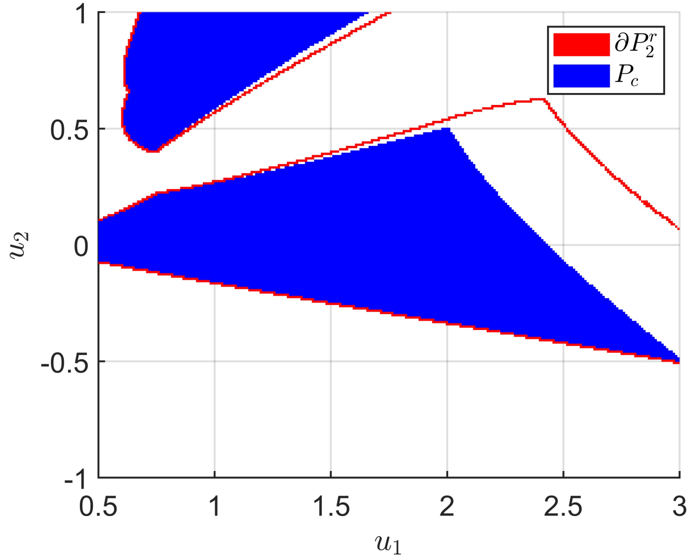
→ Computation of exact errors necessary (strong greedy algorithm, see [Haasdonk, Salomon, Wohlmuth, 2013]).

Numerical Results – Comparison of Methods

For $\varepsilon = (0.03, 0.03, 0.01, 0.01)^\top$:



(a) Strategy 1

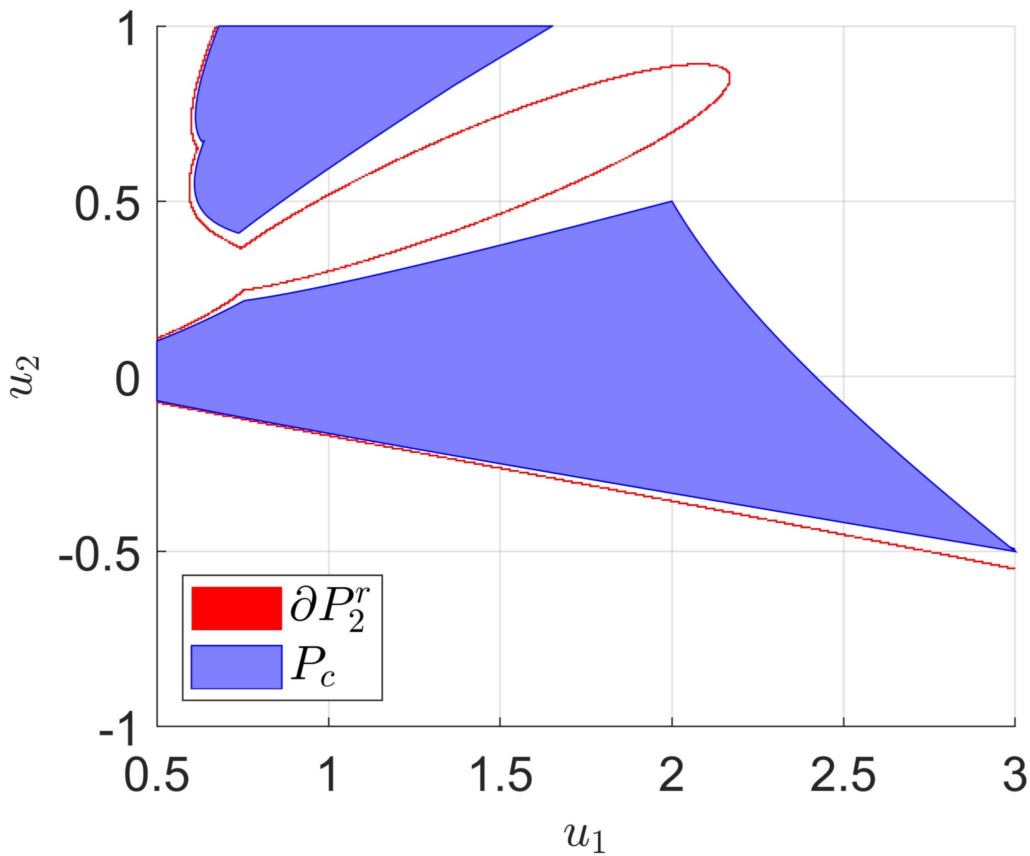


(b) Strategy 2

Algorithm	# Boxes	# Subproblems	Runtime (in seconds)
Exact cont.	15916	18746	17501s
Strategy 1	21750	24515	1426s
Strategy 2	899	1252	276s

Numerical Results – Individual Error Bounds

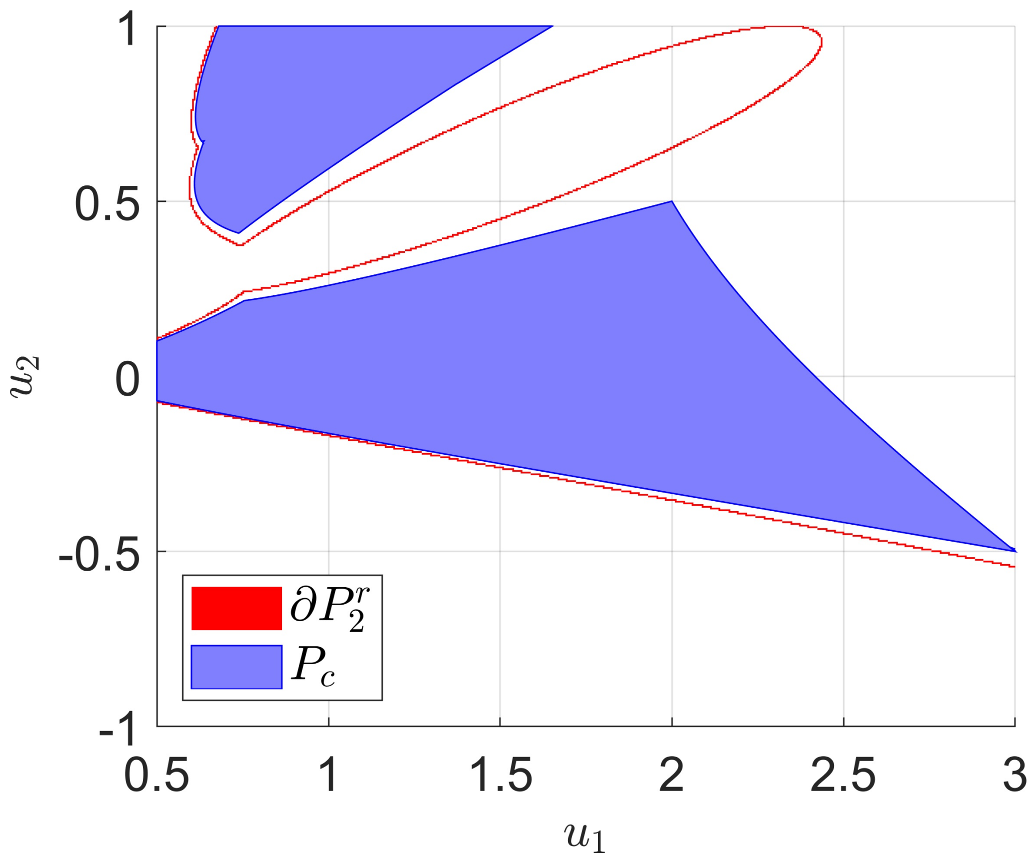
Varying the error bounds ε:



$\varepsilon = (0.1, 0.1, 0.1, 0.01)$

Numerical Results – Individual Error Bounds

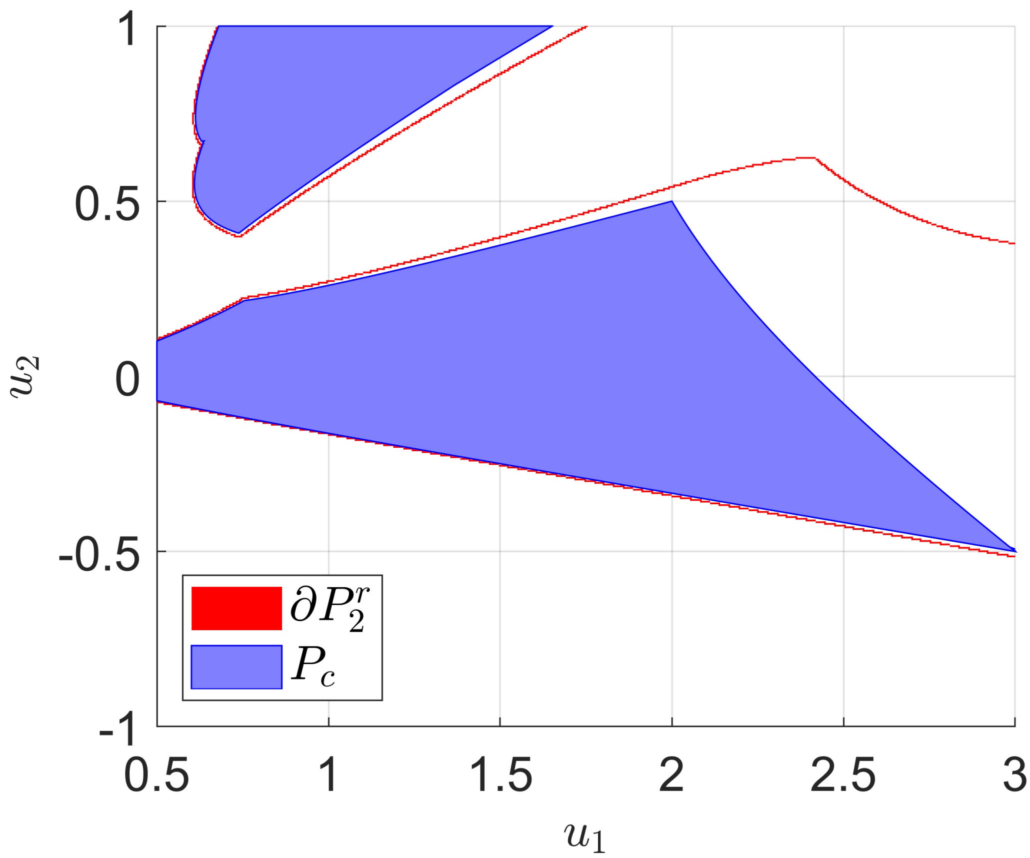
Varying the error bounds ε:



$\varepsilon = (0.0885, 0.0885, 0.0885, 0.01)$

Numerical Results – Individual Error Bounds

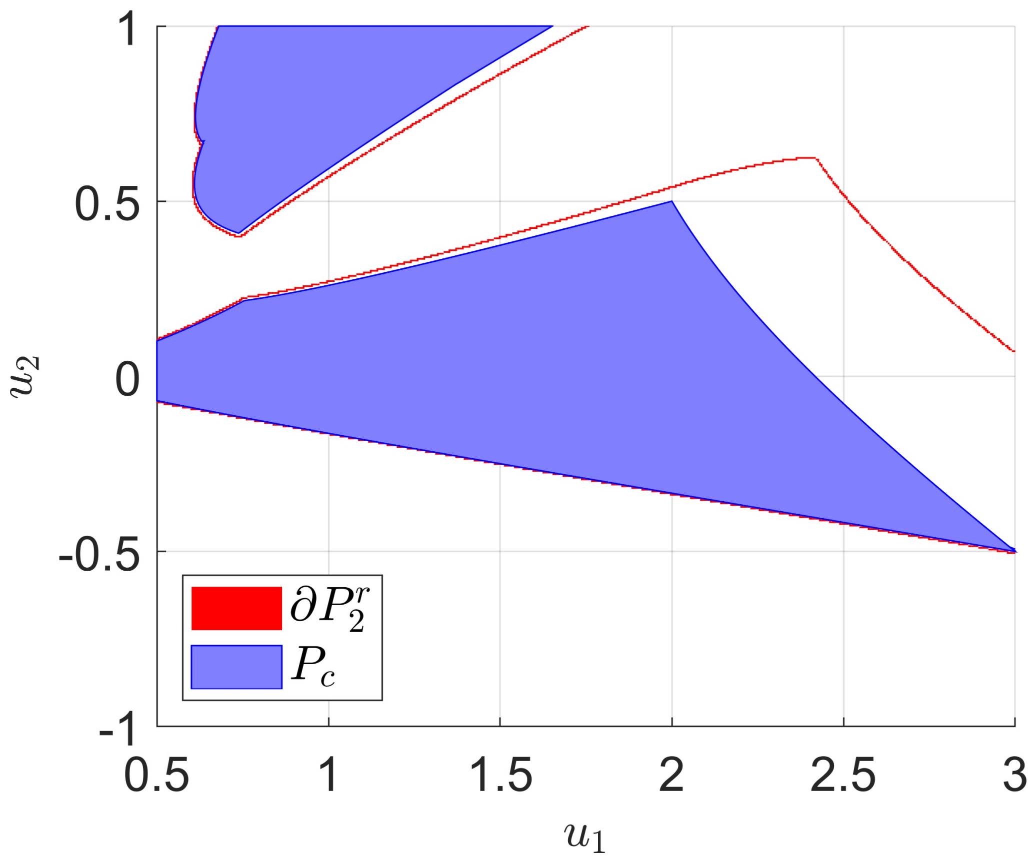
Varying the error bounds ε:



$$\varepsilon = (0.03, 0.03, 0.03, 0.01)$$

Numerical Results – Individual Error Bounds

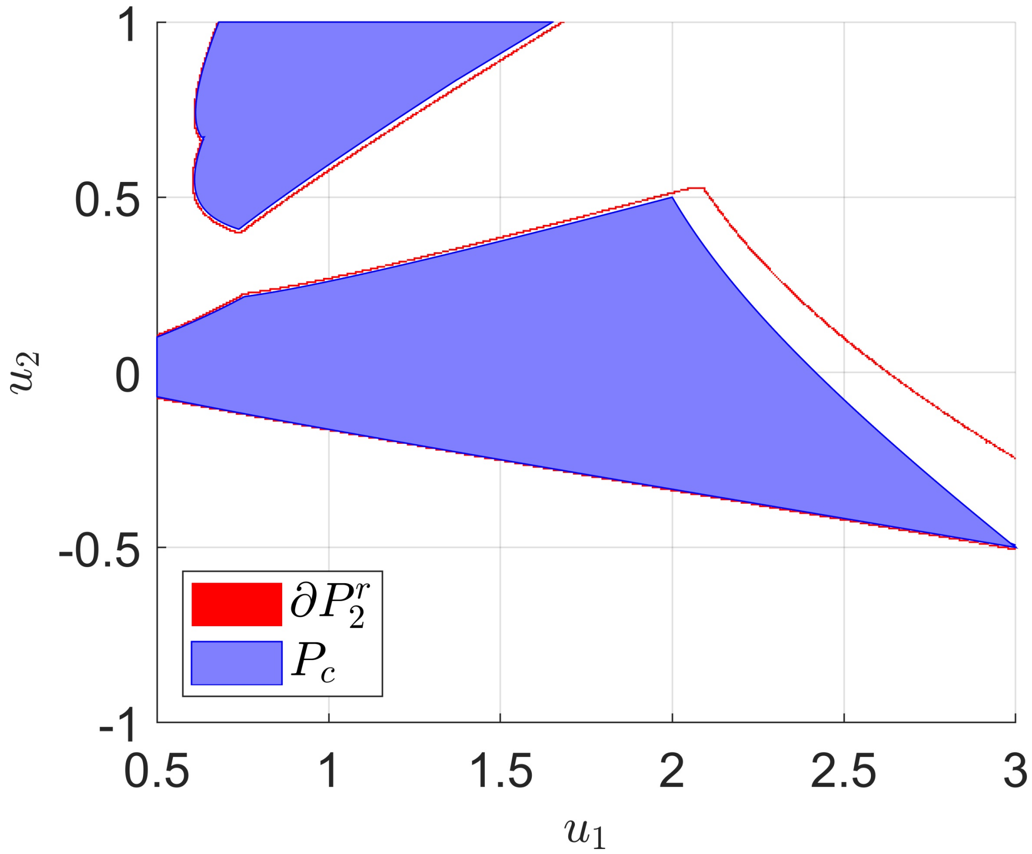
Varying the error bounds ε:



$\varepsilon = (0.03, 0.03, 0.01, 0.01)$

Numerical Results – Individual Error Bounds

Varying the error bounds ε:



$$\varepsilon = (0.03, 0.01, 0.01, 0.01)$$

Conclusion

- Solving MPOPs by a set-oriented continuation method via boxes.
- Development of efficient strategies to deal with inexactness in the gradients.
- Application of the RB method via a greedy algorithm $\rightarrow$ reduction of the computational time by a factor of up to 60.
- Individual error bounds allow a precise control of the tightness of the covering.

Conclusion

- Solving MPOPs by a set-oriented continuation method via boxes.
- Development of efficient strategies to deal with inexactness in the gradients.
- Application of the RB method via a greedy algorithm $\rightarrow$ reduction of the computational time by a factor of up to 60.
- Individual error bounds allow a precise control of the tightness of the covering.

Outlook:

- In certain situations the subproblems for computing ∂P^r are quite hard to solve $\rightarrow$ development of more sophisticated solvers
- Development of a more efficient a-posteriori error estimator for the error in the gradients.

Literature

- S. B., B. Gebken, M. Dellnitz, S. Peitz & S. Volkwein. "ROM-based multiobjective optimization of elliptic PDEs via numerical continuation". *Preprint, 2019*. https://arxiv.org/abs/1906.09075.
- B. Gebken, S. Peitz & M. Dellnitz. "On the hierarchical structure of Pareto critical sets". *Journal of Global Optimization, 2019*.
- C. Hillermeier, "Nonlinear Multiobjective Optimization: A Generalized Homotopy Approach", *Birkhäuser Basel, 2001*.
- O. Schütze, A. Dell'Aere & M. Dellnitz, "On Continuation Methods for the Numerical Treatment of Multi-Objective Optimization Problems". *Practical Approaches to Multi-Objective Optimization, 2005*.