Universität Konstanz

Reduced-Order Approaches for PDE-Constrained Multiobjective Optimization

joint work with U Dubrovnik / TU Eindhoven / U Erlangen / TU München / U Paderborn

Stefan Volkwein

University of Konstanz, Chair Numerical Optimization

New Trends in PDE-Constrained Optimization, (RICAM, Linz), October 16, 2019,

Outline of the Talk

1 Multiobjective Optimization

2 The Reference Point Method

3 ROM for Nonsmooth Optimization

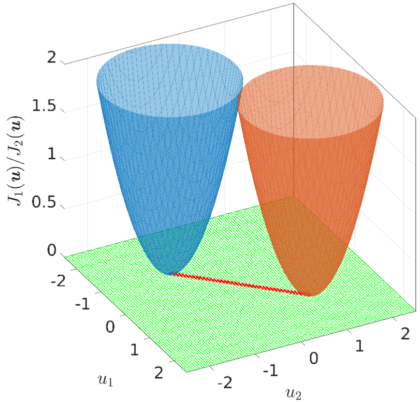
4 Greedy Controllability of Parametrized Linear Systems

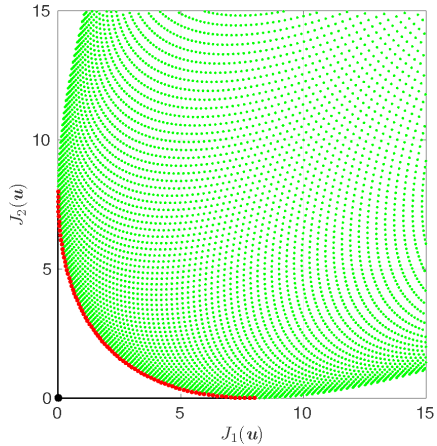
5 Conclusions

1 Multiobjective Optimization [Banholzer / Beermann / Makarov / Reichle / Spura]

UNIVERSITÄT PADERBORN

[M. Dellnitz / B. Gebken / S. Peitz]





Multiobjective / Multicriterial Optimization [e.g., Ehrgott'05]

Competing goals in many applications: e.g.

- production (quality $\longleftrightarrow$ production costs)
- transport (travel costs $\longleftrightarrow$ comfort $\longleftrightarrow$ energy consumption)

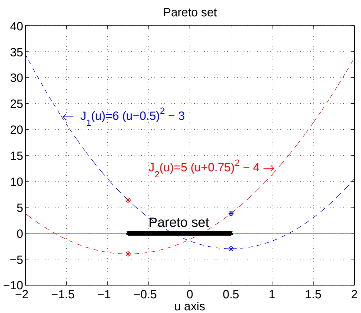
Multiobjective optimization: $\hat{J} \in C^1(\mathcal{U}_{\text{ad}}, \mathbb{R}^k)$, $\mathcal{U}$ Hilbert space

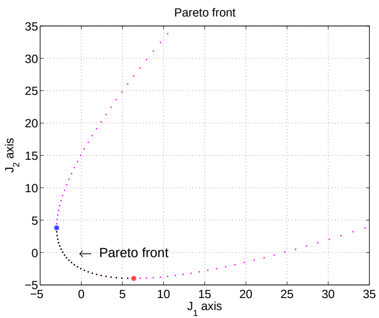
$$\min \hat{J}(u) = \big(\hat{J}_1(u), \ldots, \hat{J}_k(u)\big) \quad \text{subject to} \quad u \in \mathcal{U}_{\text{ad}} \subset \mathcal{U} \qquad (\hat{\mathbf{P}})$$

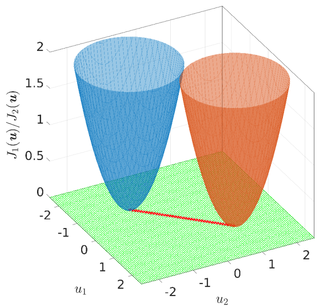
Pareto optimal points: $\bar{u} \in \mathcal{U}_{\text{ad}}$ Pareto optimal, if there is no $u \in \mathcal{U}_{\text{ad}}$ such that

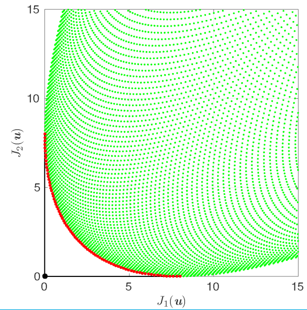
$$\hat{J}_j(u) \le \hat{J}_j(\bar{u}) \text{ for } 1 \le j \le k \quad \text{and} \quad \hat{J}_j(u) < \hat{J}_j(\bar{u}) \text{ for at least one } j \in \{1, \ldots, k\}$$

Sets: Pareto set $\mathcal{P}_{\text{s}} = \big\{\bar{u} \in \mathcal{U}_{\text{ad}} \,\big|\, \bar{u} \text{ Pareto optimal}\big\}$, Pareto front $\mathcal{P}_{\text{f}} = \hat{J}(\mathcal{P}_{\text{s}}) \subset \mathbb{R}^k$









Optimality Conditions and Optimization Methods

Karush-Kuhn-Tucker [Kuhn/Tucker'51]: $\bar{u} \in \mathcal{U}_{\text{ad}}$ Pareto optimal, then there is $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_k) \in \mathbb{R}^k$

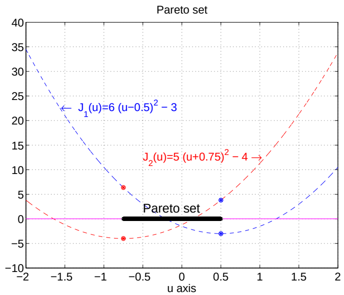
$$0 \le \bar{\mu}_i \le 1, \quad \sum_{i=1}^{k} \bar{\mu}_i = 1, \quad \sum_{i=1}^{k} \bar{\mu}_i \left\langle \hat{J}_i'(\bar{u}), u - \bar{u} \right\rangle_{\mathcal{U}} = \left\langle \sum_{i=1}^{k} \bar{\mu}_i \hat{J}_i'(\bar{u}), u - \bar{u} \right\rangle_{\mathcal{U}} \ge 0 \text{ for all } u \in \mathcal{U}_{\text{ad}}$$

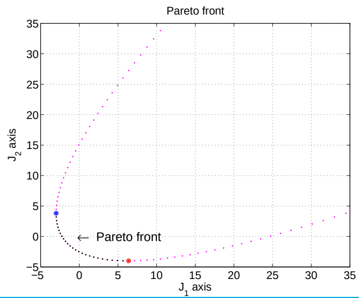
$\to$ first-order (necessary) optimality conditions for $\min\limits_{u \in \mathcal{U}_{\text{ad}}} \hat{G}(u; \bar{\mu}) = \sum_{i=1}^{k} \bar{\mu}_i \hat{J}_i(u)$

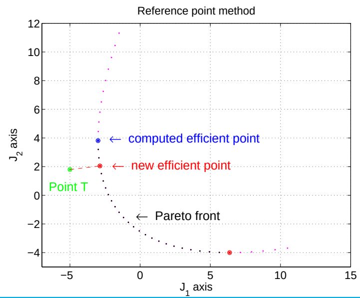
Weighted sum method: solve $\bar{u}_\mu = \operatorname*{arg\,min}\limits_{u \in \mathcal{U}_{\text{ad}}} \hat{G}(u; \mu)$ with $0 \le \mu_i \le 1$ and $\sum_{i=1}^{k} \mu_i = 1$

Euclidean reference point method [Wierzbicki'79]: $\min\limits_{u \in \mathcal{U}_{\text{ad}}} \hat{F}_T(u) = \frac{1}{2} \|T - \hat{J}(u)\|_2^2$ for reference point $T = (T_1, \ldots, T_k)^\top$

Sets: Pareto set $\mathcal{P}_{\text{s}} = \left\{ \bar{u} \in \mathcal{U}_{\text{ad}} \,\middle|\, \bar{u} \text{ Pareto optimal} \right\}$, Pareto front $\mathcal{P}_{\text{f}} = \hat{J}(\mathcal{P}_{\text{s}}) \subset \mathbb{R}^k$







2 The Reference Point Method [Banholzer / Beermann / Makarov / Reichle / Spura]

Supported by:

Federal Ministry of Economics and Technology

on the basis of a decision by the German Bundestag

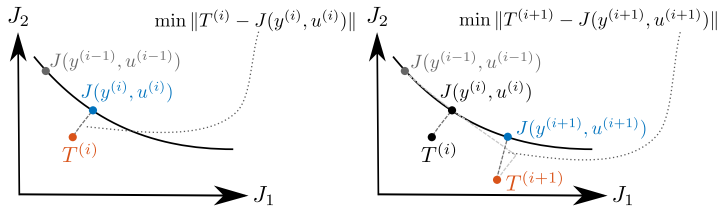
SPP 1962

DFG Deutsche Forschungsgemeinschaft

Scalarization by Introduction of a Distance Function

Ideal vector: $\hat{T} = (\hat{T}_1, \dots, \hat{T}_k)^\top \in \mathbb{R}^k$ with $\hat{T}_i = \min_{u \in \mathcal{U}_{\mathsf{ad}}} \hat{J}_i(u)$ for $i = 1, \dots, k$

Define: $g : \mathbb{R}^k \to \mathbb{R}$, where, e.g., $g(s) = \frac{1}{2}\|s\|_2^2$ or $g(s) = \|s\|_\infty$



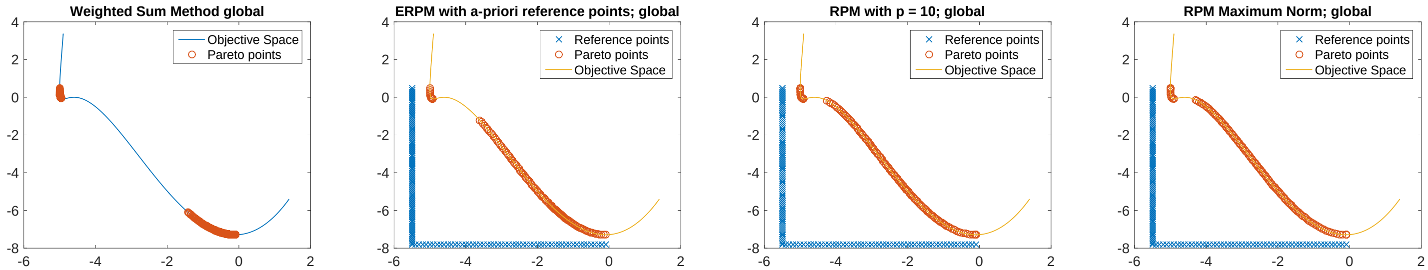
Reference point problem: for chosen $T = (T_1, \dots, T_k)^\top \in \mathbb{R}^k_{\leq \hat{T}} = \{\tilde{T} \in \mathbb{R}^k \,|\, \tilde{T} \leq \hat{T}\}$ consider

$$\min_{u \in \mathcal{U}_{\mathsf{ad}}} \hat{F}_T(u) = g\big(T - \hat{J}(u)\big) \quad \text{subject to} \quad u \in \mathcal{U}_{\mathsf{ad}} \qquad (\hat{\mathbf{P}}_T)$$

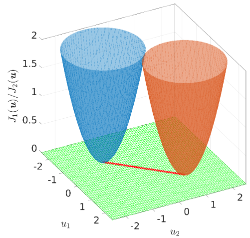
Example:

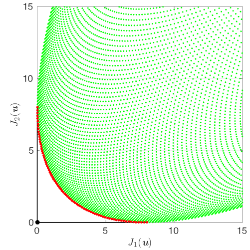
$$\hat{J}_1(u) = \frac{u^2}{10} - \frac{2u}{5} - \frac{23}{5}$$

$$\hat{J}_2(u) = \frac{u^4}{40} + \frac{2u^3}{15} - \frac{u^2}{4}$$



Euclidean Reference Point Method





Sets: Pareto set $\mathcal{P}_\mathrm{s} = \{\bar{u} \in \mathcal{U}_\mathsf{ad} \,|\, \bar{u}$ Pareto optimal$\}$

Sets: Pareto front $\mathcal{P}_\mathrm{f} = \hat{J}(\mathcal{P}_\mathrm{s}) \subset \mathbb{R}^k$

Reference point: choose $T = (T_1, \ldots, T_k)^\top \in \mathcal{P}_\mathrm{f} + \mathbb{R}^k_{\leq 0} = \{z + x \,|\, z \in \mathcal{P}_\mathrm{f}$ and $\mathbb{R}^k \ni x \leq 0\}$

Reference point problem: consider

$$\min_{u \in \mathcal{U}_\mathsf{ad}} \hat{F}_T(u) = \frac{1}{2} \left\| T - \hat{J}(u) \right\|_2^2 = \frac{1}{2} \sum_{i=1}^k \left| T_i - \hat{J}_i(u) \right|^2 \quad \text{subject to} \quad u \in \mathcal{U}_\mathsf{ad} \tag{$\hat{\mathbf{P}}_T$}$$

First derivative: $\hat{F}_T'(u) = \sum_{i=1}^k (\hat{J}_i(u) - T_i) \langle \hat{J}_i'(u), \bullet \rangle_\mathcal{U} \in \mathscr{L}(\mathcal{U}, \mathbb{R})$

First-order optimality conditions: If $\bar{u}_T$ solves $(\hat{\mathbf{P}}_T)$ we have

$$\sum_{i=1}^k \left(\hat{J}_i(\bar{u}_T) - T_i \right) \left\langle \hat{J}_i'(\bar{u}_T), u - \bar{u}_T \right\rangle_\mathcal{U} \geq 0 \quad \text{for all } u \in \mathcal{U}_\mathsf{ad}$$

Computation of the Pareto Set and Pareto Front

Sets: Pareto set $\mathcal{P}_s = \{\bar{u} \in \mathcal{U}_{ad} \,|\, \bar{u} \text{ Pareto optimal}\}$, Pareto front $\mathcal{P}_f = \hat{J}(\mathcal{P}) \subset \mathbb{R}^k$

Algorithm 1 (Euclidean reference point method)

1: Choose a reference point $T^{(1)} = (T_1^{(1)}, \ldots, T_k^{(1)})^\top$ and set $i = 1$;
2: Compute for the distance function

$$\hat{F}^{(i)}(u) = \frac{1}{2}\left\|T^{(i)} - \hat{J}(u)\right\|_2^2 = \frac{1}{2}\sum_{j=1}^{k}\left|T_j^{(i)} - \hat{J}_j(u)\right|^2 \geq 0$$

the solution to the scalar reference point problem

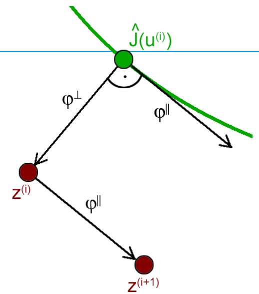
$$\bar{u}_T^{(i)} = \arg\min\left\{\hat{F}_T^{(i)}(u) \,\middle|\, u \in \mathcal{U}_{ad}\right\} \qquad (\hat{\mathbf{P}}_T^{(i)})$$

3: Choose a new reference point $T^{(i+1)}$ and go back to step 2.

Variation of the reference points: for $\hat{J}^i = \hat{J}(\bar{u}_T^{(i)})$ set

$$T^{(i+1)} := \hat{J}^i + h^{||} \cdot \frac{\varphi^{||}}{\|\varphi^{||}\|} + h^{\perp} \cdot \frac{\varphi^{\perp}}{\|\varphi^{\perp}\|} \quad \text{for } i \geq 1$$

and choose $h^{||}, h^{\perp} > 0$ to control the coarseness of $\mathcal{P}_f$ points



Heat Equation with Convection

Bicriterial optimal control problem: minimize

$$J(y,u) = \frac{1}{2}\begin{pmatrix} \int_0^{t_\mathsf{f}}\int_\Omega \big|y(t,\boldsymbol{x}) - y^\mathsf{d}(t,\boldsymbol{x})\big|^2\,\mathrm{d}\boldsymbol{x}\mathrm{d}t \\ \sum_{i=1}^m \int_0^{t_\mathsf{f}} \big|u_i(t)\big|^2\,\mathrm{d}t \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \|y-y^\mathsf{d}\|^2_{L^2(Q)} \\ \|u\|^2_{\mathcal{U}} \end{pmatrix}$$

subject to the parabolic convection-diffusion PDE

$$\begin{aligned} y_t(t,\boldsymbol{x}) - \Delta y(t,\boldsymbol{x}) + b(t,\boldsymbol{x})\cdot\nabla y(t,\boldsymbol{x}) &= \sum_{i=1}^m u_i(t)\chi_{\Omega_i}(\boldsymbol{x}), && (t,\boldsymbol{x})\in Q=(0,t_\mathsf{f})\times\Omega,\ \Omega=\Omega_1\dot\cup\ldots\dot\cup\Omega_m \\ \frac{\partial y}{\partial \boldsymbol{n}}(t,\boldsymbol{x}) + \alpha_i y(t,\boldsymbol{x}) &= \alpha_i y_\mathsf{a}(t), && (t,\boldsymbol{x})\in\Sigma_i=(0,t_\mathsf{f})\times\Gamma_i,\ 1\le i\le r \\ y(0,\boldsymbol{x}) &= y_\circ(\boldsymbol{x}), && \boldsymbol{x}\in\Omega\subset\mathbb{R}^d,\ d\in\{1,2,3\} \end{aligned} \qquad \textbf{(SE)}$$

and the bilateral control constraints $u\in\mathcal{U}_\mathsf{ad} = \big\{u\in\mathcal{U}=L^2(0,t_\mathsf{f};\mathbb{R}^m)\,\big|\,u_\mathsf{a}(t)\le u(t)\le u_\mathsf{b}(t)\text{ in }[0,t_\mathsf{f}]\big\}$

Assumptions: $y^\mathsf{d}\in L^2(0,t_\mathsf{f};L^2(\Omega))$, b bounded, $\chi_{\Omega_i}\in L^2(\Omega)$, $\alpha_i\ge 0$, $y_\mathsf{a}\in L^2(0,t_\mathsf{f})$, $y_\circ\in L^2(\Omega)$, $u_\mathsf{a}\le u_\mathsf{b}$ in $\mathcal{U}$

Bilinear form for (SE): for $\varphi,\psi\in H^1(\Omega)$ and $t\in[0,t_\mathsf{f}]$ define

$$a(t;\varphi,\psi) = \int_\Omega \nabla\varphi\cdot\nabla\psi + \big(b(t,\cdot)\nabla\varphi\big)\psi\,\mathrm{d}\boldsymbol{x} + \sum_{i=1}^r \alpha_i\int_{\Gamma_i}\varphi\psi\,\mathrm{d}\boldsymbol{x}, \quad \langle g_\mathsf{a}(t),\varphi\rangle = \sum_{i=1}^r \alpha_i\int_{\Gamma_i} y_\mathsf{a}(t)\varphi\,\mathrm{d}\boldsymbol{x}$$

Reduced Formulation for the Multiobjective Problem

Bilinear form for (SE): for $\varphi, \psi \in H^1(\Omega)$ and $t \in [0, t_\mathsf{f}]$ define

$$a(t;\varphi,\psi) = \int_\Omega \nabla\varphi \cdot \nabla\psi + \big(b(t,\cdot)\cdot\nabla\varphi\big)\psi\,\mathrm{d}\boldsymbol{x} + \sum_{i=1}^r \alpha_i \int_{\Gamma_i} \varphi\psi\,\mathrm{d}\boldsymbol{x}, \quad \langle g_\mathsf{a}(t),\varphi\rangle = \sum_{i=1}^r \alpha_i \int_{\Gamma_i} y_\mathsf{a}(t)\varphi\,\mathrm{d}\boldsymbol{x}$$

Control-to-state operator: $y = \mathscr{S}(u)$ solves for given control $u \in \mathfrak{U} = L^2(0,t_\mathsf{f};\mathbb{R}^m)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle y(t),\varphi\rangle_{L^2(\Omega)} + a(t;y(t),\varphi) = \langle g_\mathsf{a}(t),\varphi\rangle + \sum_{i=1}^m u_i(t)\int_{\Omega_i}\varphi\,\mathrm{d}\boldsymbol{x} \text{ for all } \varphi \in H^1(\Omega) \quad \text{and} \quad y(0) = y_\circ$$

Reduced cost functional: $\hat{J}(u) = \begin{pmatrix} \hat{J}_1(u) \\ \hat{J}_2(u) \end{pmatrix} = \begin{pmatrix} J_1(\mathscr{S}(u),u) \\ J_2(\mathscr{S}(u),u) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \|\mathscr{S}(u) - y^\mathsf{d}\|^2_{L^2(Q)} \\ \|u\|^2_{\mathfrak{U}} \end{pmatrix}$ for $u \in \mathfrak{U}$

Reduced multicriterial optimal control problem:

$$\min \hat{J}(u) \quad \text{subject to} \quad u \in \mathfrak{U}_\mathsf{ad} = \big\{u \in \mathfrak{U}\,\big|\, u_\mathsf{a} \le u \le u_\mathsf{b} \text{ in } [0,t_\mathsf{f}]\big\} \tag{$\hat{\mathbf{P}}$}$$

A-Posteriori Error Estimation

Euclidean reference point problem: for $T = (T_1, T_2) \in \mathbb{R}^2$ consider

$$\min \hat{F}_T(u) = \frac{1}{2}\|T - \hat{J}(u)\|_2^2 = \frac{1}{2}\big(T_1 - \hat{J}_1(u)\big)^2 + \frac{1}{2}\big(T_2 - \hat{J}_2(u)\big)^2 \quad \text{subject to} \quad u \in \mathfrak{U}_{\mathsf{ad}} \qquad (\hat{\mathbf{P}}_T)$$

First-order sufficient optimality condition for $(\hat{\mathbf{P}}_T)$: optimal $\bar{u}_T \in \mathfrak{U}$ satisfies

$$\langle \hat{F}_T'(\bar{u}_T), u - \bar{u}_T\rangle_{\mathfrak{U}} = \sum_{i=1}^{2} (\hat{J}_i(u) - T_i)\langle \hat{J}_i'(u), u - \bar{u}_T\rangle_{\mathfrak{U}} \geq 0 \quad \text{for all } u \in \mathfrak{U}_{\mathsf{ad}} \qquad (1)$$

Second-order sufficient optimality condition for $(\hat{\mathbf{P}}_T)$: For any $u \in \mathfrak{U}_{\mathsf{ad}}$ with $\hat{J}_1(u) \geq T_1$ the hessian $\hat{F}_T''(u)$ satisfies

$$\langle \hat{F}_T''(u)\tilde{u}, \tilde{u}\rangle_{\mathfrak{U}} \geq \big(\hat{J}_2(u) - T_2\big)\,\|\tilde{u}\|_{\mathfrak{U}}^2 \quad \text{for all } \tilde{u} \in \mathfrak{U}$$

If $T_2 < \hat{T}_2 = \min_{\tilde{u}\in\mathfrak{U}_{\mathsf{ad}}} \hat{J}_2(\tilde{u})$: $\langle \hat{F}_T''(u)\tilde{u}, \tilde{u}\rangle_{\mathfrak{U}} \geq \bar{\kappa}_T\,\|\tilde{u}\|_{\mathfrak{U}}^2$ for all $\tilde{u} \in \mathfrak{U}$ with $\bar{\kappa}_T = \min\limits_{u\in\mathfrak{U}_{\mathsf{ad}}} \hat{J}_2(u) - T_2 = \hat{T}_2 - T_2 > 0$

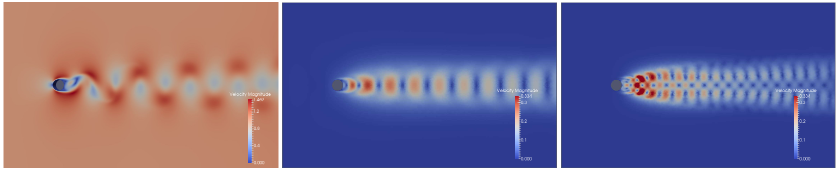
A-posteriori error estimate: for any $\bar{u}_T^{\mathsf{apo}} \in \mathfrak{U}_{\mathsf{ad}}$ there is a computable „perturbation" $\zeta_T^{\mathsf{apo}} \in \mathfrak{U}$ satisfying

$$\|\bar{u}_T - \bar{u}_T^{\mathsf{apo}}\|_{\mathfrak{U}} \leq \Delta_T^{\mathsf{apo}} \quad \text{with } \Delta_T^{\mathsf{apo}} = \frac{1}{\bar{\kappa}_T}\|\zeta_T^{\mathsf{apo}}\|_{\mathfrak{U}} \qquad \longrightarrow \qquad \text{apply } \textbf{Reduced-Order Modeling (ROM)}$$

Reduced-Order Modeling (ROM) using Proper Orthogonal Decomposition (POD)

Scalar reference point problem:

$$\min \hat{F}_T(u) = \tfrac{1}{2}\left(T_1 - \hat{J}_1(u)\right)^2 + \tfrac{1}{2}\left(T_2 - \hat{J}_2(u)\right)^2 \text{ s.t. } u \in \mathcal{U}_{\mathrm{ad}} \qquad (\hat{\mathbf{P}}_T)$$



Discretization: finite elements (FE) for x, implicit Euler for t

ROM: apply POD model reduction for the spatial discretization

Algorithm 2 (POD method)

1: Determine $y^{\mathrm{FE}} = \mathscr{S}^{\mathrm{FE}}(\bar{u}_T)$ for a computed solution to $(\hat{\mathbf{P}}_T)$ and solve (with, e.g., trapezoidal) weights $\alpha_j > 0$

$$\min \sum_{j=1}^{n} \alpha_j \left\| y^{\mathrm{FE}}(t_j) - \sum_{i=1}^{\ell} \langle y^{\mathrm{FE}}(t_j), \psi_i \rangle \, \psi_i \right\|^2 \quad \text{s.t.} \quad \{\psi_i\}_{i=1}^{\ell} \text{ are orthonormal} \qquad (\ell \ll \#\text{ FE degrees of freedom})$$

$\longrightarrow \{\psi_i\}_{i=1}^{\ell}$ given as a solution of an eigenvalue problem and

$$\sum_{j=1}^{n} \alpha_j \left\| y^{\mathrm{FE}}(t_j) - \sum_{i=1}^{\ell} \langle y^{\mathrm{FE}}(t_j), \psi_i \rangle \, \psi_i \right\|^2 = \sum_{i > \ell+1} \lambda_i, \quad \lambda_1 \geq \lambda_2 \geq \ldots \geq 0$$

2: Use POD basis in a spatial Galerkin discretization of the cost: $\hat{J}^{\ell}(u) = J(\mathscr{S}^{\ell}(u), u)$;

3: Solve

$$\min \hat{F}_T^{\ell}(u) = \tfrac{1}{2}\left(T_1 - \hat{J}_1^{\ell}(u)\right)^2 + \tfrac{1}{2}\left(T_2 - \hat{J}_2^{\ell}(u)\right)^2 \quad \text{s.t.} \quad u \in \mathcal{U}_{\mathrm{ad}} \qquad (\hat{\mathbf{P}}_T^{\ell})$$

to get POD suboptimal control $\bar{u}_T^{\ell}$ which satisfies $\|\bar{u}_T^{\ell} - \bar{u}_T\|_{\mathcal{U}} \leq \Delta_T^{\mathsf{apo}}$ with $\Delta_T^{\mathsf{apo}} = \|\zeta_T^{\mathsf{apo}}\|_{\mathcal{U}} / \bar{\kappa}_T$ and computable ζ_T^{apo}, $\bar{\kappa}_T$.

How can we ensure a small error?

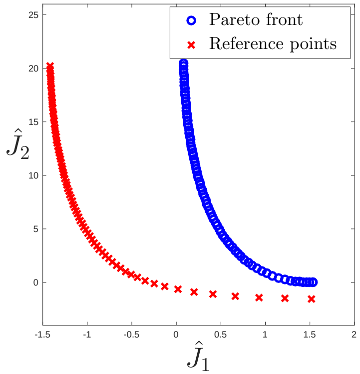
A-posteriori error estimate: $\|\bar{u}_T^\ell - \bar{u}_T\|_{\mathcal{U}} \le \Delta_T^{\mathsf{apo}}$ with $\Delta_T^{\mathsf{apo}} = \|\zeta_T^{\mathsf{apo}}\|_{\mathcal{U}} / \bar{\kappa}_T$ and $\zeta_T^{\mathsf{apo}} = \zeta_T^{\mathsf{apo}}(\bar{u}_T^\ell)$

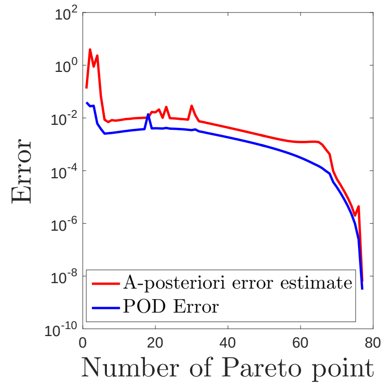
Algorithm 3 (POD-based reference point method)

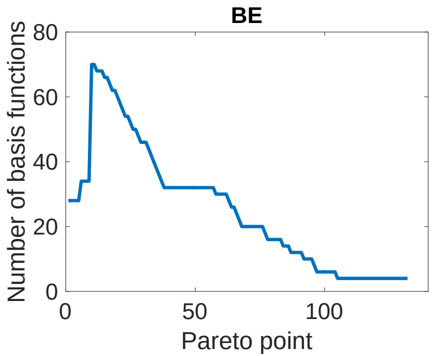
Require: $\tau > 0$ (upper bound on the a-posterior error estimate) and $0 < \eta \ll 1$ (scalar factor);
1: Compute the minimizers of $\hat{J}_1$ and $\hat{J}_2$;
2: Compute a POD basis with the optimal states and adjoints of the minimizers;
3: **while** *end of Pareto front is not reached* **do**
4: Generate a new reference point;
5: Compute solution $\bar{u}_T^\ell$ to $(\hat{\mathbf{P}}_T^\ell)$;
6: Compute the a-posteriori error estimate Δ_T^{apo};
7: **if** $\Delta_T^{\mathsf{apo}} < \tau$ **then**
8: **if** $\Delta_T^{\mathsf{apo}} < \eta\tau$ **then**
9: Reduce the number ℓ of POD basis functions;
10: **end if**
11: **else**
12: Extend (e.g., $\ell = \ell + 1$) or even update POD basis and go back to 5;
13: **end if**
14: **end while**

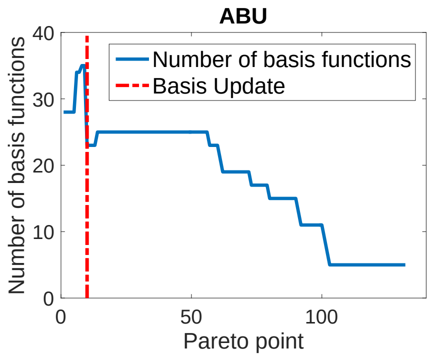
Numerical Test 1: Two-Dimensional Example [Banholzer/Makarov/V.'18]

Speed-ups	
Basis extension (BE)	14
Adaptive basis update ABU)	18



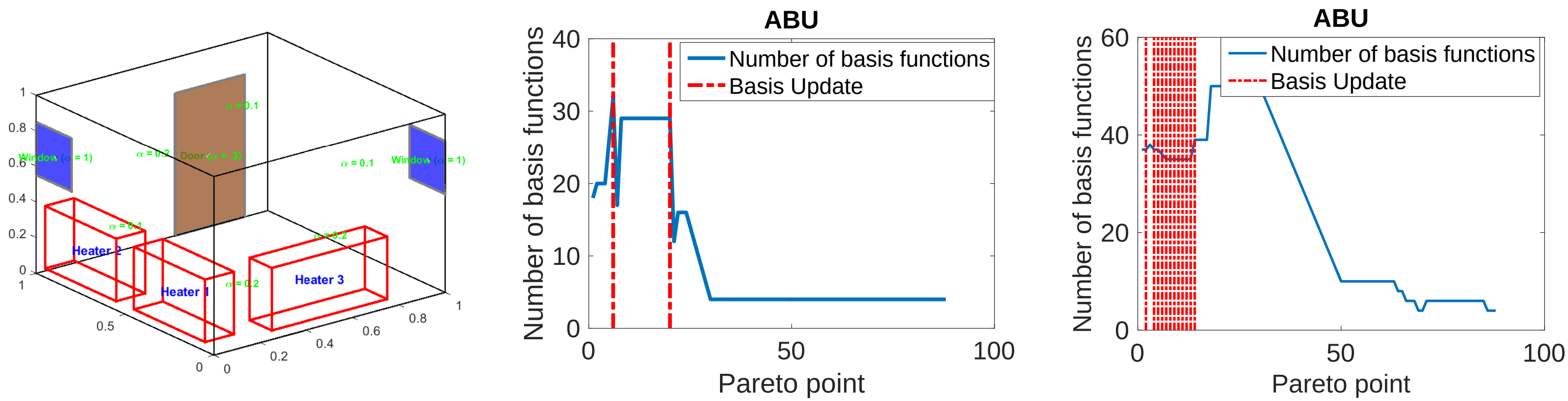






Numerical Test 2: Three-Dimensional Example [Banholzer / Spura / V.'18]

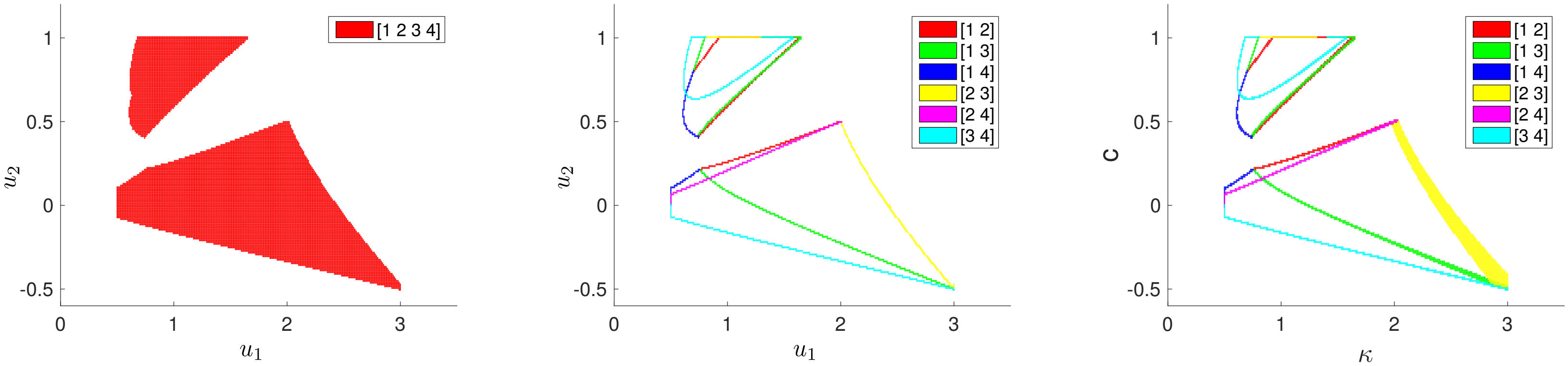
PDE constraint: $y_t(t,x) - \kappa\Delta y(t,x) + c_b b(t,x)\cdot\nabla y(t,x) = \sum_{i=1}^{3} u_i(t)\chi_i(x)$ in Q



Figures: geometry Ω / $c_b = 1$ / $c_b = 4$

Extension to More Than Two Objectives

Problem: $\min \hat{J}(u) = \frac{1}{2}\begin{pmatrix} \|y(u)-y_1^{\mathsf{d}}\|^2_{L^2(\Omega)} \\ \|y(u)-y_2^{\mathsf{d}}\|^2_{L^2(\Omega)} \\ \|y(u)-y_3^{\mathsf{d}}\|^2_{L^2(\Omega)} \\ u_1^2+u_2^2 \end{pmatrix}$ s.t. $u \in \mathcal{U}_{\mathsf{ad}} \subset \mathbb{R}^2$ and $-u_1\Delta y(x) + u_2 b(x)\cdot\nabla y(x) + 0.5y(x) = f(x)$ in Ω + Robin-BC



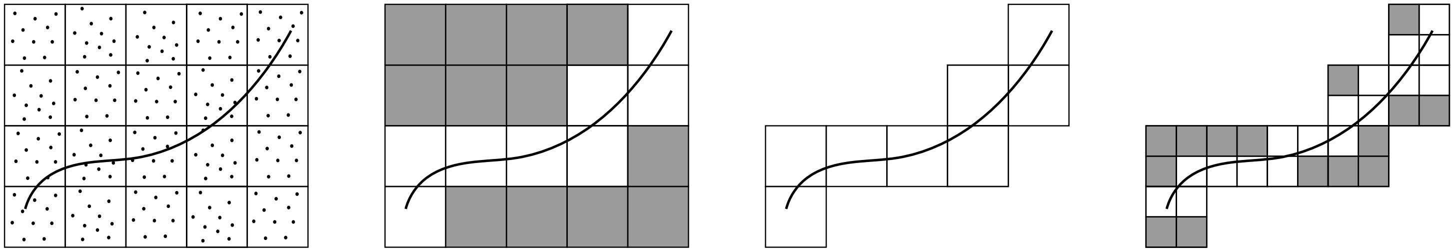
Left plot: Pareto set (FEM discretization) with two parameters and four objectives.

Middle plot: The six Pareto sets considering only two objectives can be computed significantly faster.

Right plot: Approximation via RB-based ROMs with a certified error bound (RB – Reduced Basis ... explained later)

Set-Oriented Methods – Subdivision Algorithm

Global, gradient free subdivision algorithm:



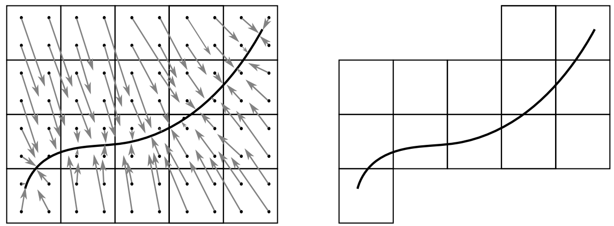
Sampling   Nondominance test   Eliminate dominated boxes   Next iteration

Gradient-based subdivision algorithm:

– Compute $\hat{\alpha} = \arg\min \Big\{ \big\| \sum_{i=1}^{k} \alpha_i \hat{J}_i'(u) \big\|_2^2 \;\Big|\; \alpha \in \mathbb{R}^k,\ \sum_{i=1}^{k} \alpha_i = 1,\ \alpha_1, \ldots, \alpha_k \geq 0 \Big\}$.

$\rightarrow$ $q(u) = -\sum_{i=1}^{k} \hat{\alpha}_i \hat{J}_i'(u)$ is a descent direction or $q(u) = 0$

$\rightarrow$ ODE $\dot{u}(t) = q(u(t))$ has an attracting set containing all stationary points u

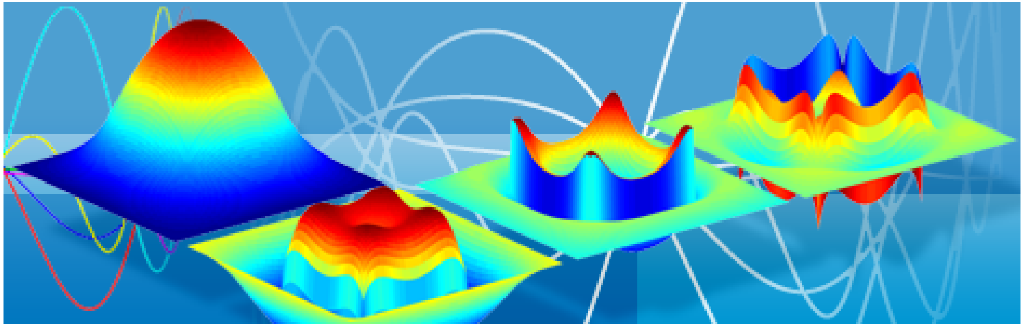


Sample gradients   Eliminate boxes

Continuation method: $\begin{pmatrix} \sum_{i=1}^{k} \alpha_i \hat{J}_i'(u) \\ 1 - \sum_{i=1}^{k} \alpha_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and inexact gradients

$\rightarrow$ talk of **Stefan Banholzer** on Thursday (11:00–11:30)

3 ROM for Nonsmooth Optimization [Bernreuther / Müller]



DFG Deutsche Forschungsgemeinschaft

SPP 1962

Reduced Basis (RB) Method for the Max-PDE

Consider the parametrized max-PDE: $-\Delta y(\boldsymbol{x}) + 8\pi\mu \max\{0, y(\boldsymbol{x})\} = f(\boldsymbol{x};\mu)$ in $\mathcal{D} \subset \mathbb{R}^2$, $\mu \in \mathcal{M} = [0,3]$ and $y(\boldsymbol{x}) = 0$ on $\mathcal{D}$

Finite element (FE) Galerkin formulation:

$$\int_{\mathcal{D}} \nabla y^h \cdot \nabla \varphi^h + 8\pi\mu \max\{0, y^h\} \varphi^h \,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{D}} f(\cdot\,;\mu)\varphi^h \,\mathrm{d}\boldsymbol{x} \quad \text{for all } \varphi^h \in V^h = \operatorname{span}\{\varphi_1, \dots, \varphi_m\} \subset V = H_0^1(\mathcal{D})$$

Reduced basis (RB) approach: Find $\mathscr{M}_\ell = \{\mu_1, \dots, \mu_\ell\} \subset \mathcal{M}$ such that $y^h(\mu) \approx y^\ell(\mu) = \sum_{i=1}^{\ell} \alpha_i(\mu) y^h(\mu_i)$

RB Galerkin formulation: $V^\ell = \operatorname{span}\{\psi_1, \dots, \psi_\ell\} \subset V^h$ with $\{\psi_i\}_{i=1}^\ell := \operatorname{orth}\{y^h(\mu_1), \dots, y^h(\mu_\ell)\}$

$$\int_{\mathcal{D}} \nabla y^\ell \cdot \nabla \psi + 8\pi\mu \max\{0, y^\ell\} \psi \,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{D}} f(\cdot\,;\mu)\psi \,\mathrm{d}\boldsymbol{x} \quad \text{for all } \psi \in V^\ell$$

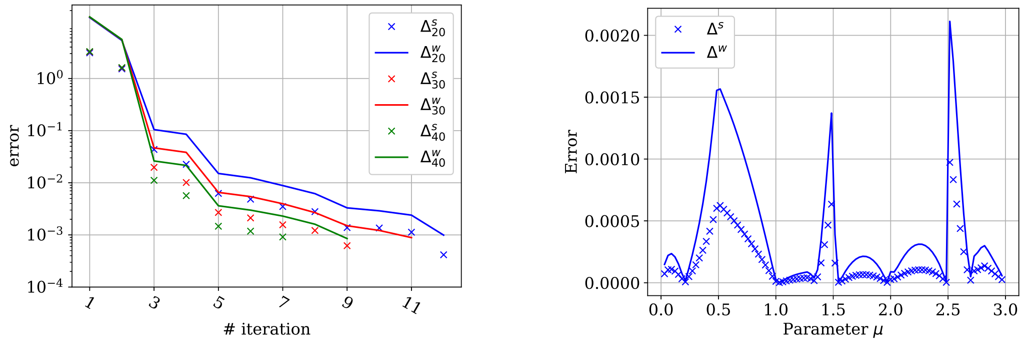
Error estimate: $\|y^h(\mu) - y^\ell(\mu)\|_V \leq \Delta^\ell_{\mathsf{apo}}(\mu) := c\,\|\Delta y^\ell - 8\pi\mu \max\{0, y^\ell\} + f(\cdot\,;\mu)\|_{(V^h)'}$ for every $\mu \in \mathcal{M}$

Greedy basis updates: Improve approximation by

$$\mu_{\ell+1} \in \operatorname*{arg\,max}_{\mu \in \mathcal{S}_{\mathsf{train}}} \begin{Bmatrix} \Delta^s = \|y^h(\mu) - y^\ell(\mu)\|_V & \text{(strong)} \\ \Delta^w = \Delta^\ell_{\mathsf{apo}}(\mu) & \text{(weak)} \end{Bmatrix}, \qquad \mathcal{M}^{\ell+1} = \mathcal{M}^\ell \cup \{\mu_{\ell+1}\}, \qquad \{\psi_i\}_{i=1}^{\ell+1} = \operatorname{orth}\Big(\{\psi_i\}_{i=1}^{\ell}, y^h(\mu_{\ell+1})\Big)$$

Numerical Results for the Max-PDE

Figure: Error estimators $\Delta^s = \|y^h(\mu) - y^\ell(\mu)\|_V$ (strong greedy) and $\Delta^w = \Delta^\ell_{\text{apo}}(\mu)$ offline (left) and online (right)



Speed-ups by the RB method compared to FE:

$n_{x_1} \times n_{x_2}$	RB with Δ^s	RB with Δ^w	RB-DEIM with Δ^s	RB-DEIM with Δ^w
20×20	5.1	5.0	6.5	6.5
30×30	16.1	13.4	23.2	19.2
40×40	39.9	32.1	60.5	47.6

Extensions

Two-dimensional parameter space: even larger speed-ups

Scalar optimization:

– good approximation results
– large speed-ups
– mesh independence for the ROM-based semismooth Newton method

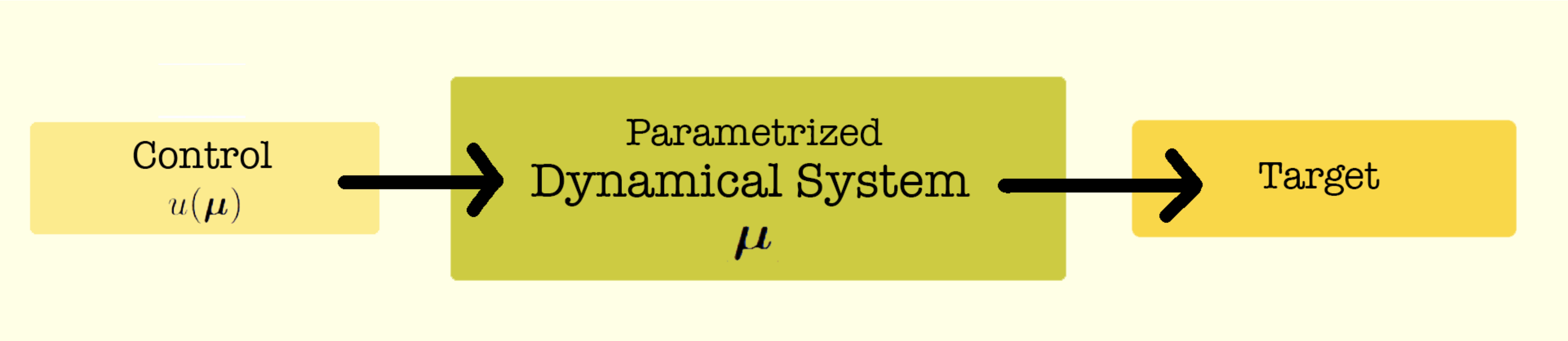
Multiobjective optimization: first results by Christof / Müller for

$$\min J(y,u) = \begin{pmatrix} j_1(y) + \frac{\sigma_1}{2}\|u\|_{L^2}^2 \\ \vdots \\ j_k(y) + \frac{\sigma_k}{2}\|u\|_{L^2}^2 \end{pmatrix} \quad \text{s.t.} \quad -\Delta y + \max\{0, y\} = u \text{ in } \mathcal{D},\ y = 0 \text{ on } \partial\mathcal{D}$$

with $\sigma_1, \ldots, \sigma_{k-1} \geq 0$ and $\sigma_k > 0$

$\rightarrow$ more at **Workshop 4: Nonsmooth optimization** in November 2019

4 Greedy Controllability [Fabrini / Iapichino / Lazar / Zuazua]



M. Lazar / E. Zuazua: *Greedy controllability of finite dimensional linear systems*, Automatica, 2016

Problem Formulation

Finite-dimensional linear control system (e.g., after spatial discretization of a PDE):

$$\dot{x}(t) = \mathrm{A}_\mu x(t) + \mathrm{B}_\mu u(t) \text{ for } t \in (0, t_\mathsf{f}] \quad \text{and} \quad x(0) = x^\circ \qquad (\mathbf{DS}_\mu)$$

State variable: $x = x_\mu : [0, t_\mathsf{f}] \to \mathbb{R}^n$ with large n

Control variable: $u = u_\mu : [0, t_\mathsf{f}] \to \mathbb{R}^m$ with $m \leq n$

Parameter set: $\mathcal{D}_\mathsf{ad} = \{\mu \in \mathbb{R}^\wp \,|\, \mu_\mathsf{a} \leq \mu \leq \mu_\mathsf{b} \text{ in } \mathbb{R}^\wp\}$, i.e., compact

System matrices: $\mathcal{D}_\mathsf{ad} \ni \mu \mapsto \mathrm{A}_\mu \in \mathbb{R}^{n \times n}$ and $\mathcal{D}_\mathsf{ad} \ni \mu \mapsto \mathrm{B}_\mu \in \mathbb{R}^{n \times m}$ Lipschitz-continuous

Controllability problem for every $\mu \in \mathcal{D}_{ad}$

Given a final target $x^1 \in \mathbb{R}^n$ find a control $u = u_\mu$ such that the state $x = x_\mu$ of $(\mathbf{DS}_\mu)$ satisfies

$$x_\mu(t_\mathsf{f}) = x^1 \quad \text{for every } \mu \in \mathcal{D}_\mathsf{ad}$$

Assumption: System $(\mathbf{DS}_\mu)$ is controllable for all values of $\mu \in \mathcal{D}_\mathsf{ad}$.

Solution of the Controllability Problem

Solution approach: If $(\varphi_\mu, \varphi_\mu^\circ)$ solves the linear-quadratic problem

$$\begin{cases} \min J_\mu(\varphi, \varphi^\circ) = \frac{1}{2} \int_0^{t_f} \|B_\mu^\top \varphi(t)\|^2 \, dt + \langle x^1, \varphi^\circ \rangle + \langle x^\circ, \varphi(0) \rangle \\ \text{s.t. } \dot{\varphi}(t) = -A_\mu^\top \varphi(t) \text{ for } t \in [0, t_f) \quad \text{and} \quad \varphi(t_f) = \varphi^\circ \end{cases}$$

then $u_\mu = B_\mu^\top \varphi_\mu$ is the control that steers the solution $x = x_\mu$ of

$$\dot{x}(t) = A_\mu x(t) + B_\mu u_\mu(t) \text{ for } t \in (0, t_f] \quad \text{and} \quad x(0) = x^\circ$$

to the desired target x^1; see, e.g., [Mica/Zuazua'05]

Numerical strategy: apply CG method to minimize $\hat{J}_\mu(\varphi^\circ) = J_\mu(\varphi_\mu(\varphi^\circ), \varphi^\circ)$

Problem: computationally too expensive in the multi-query context

Goal [Lazar/Zuazua'16]

For $\varepsilon > 0$ find *offline* $N(\varepsilon)$ parameters μ_j and associated controls $u_j = u_{\mu_j}$, so that $u_\mu = \sum_{j=1}^{N(\varepsilon)} \alpha_\mu^j u_j$ can be rapidly computed *online* with $\|x_\mu(t_f) - x^1\| \leq \varepsilon$ for any $\mu \in \mathcal{D}_{ad}$.

Main Idea

Goal [Lazar/Zuazua'16]

Given a tolerance $\varepsilon > 0$ find $N(\varepsilon)$ parameters $\mu_1, \ldots, \mu_{N(\varepsilon)}$, so that for all $\mu \in \mathcal{D}_{\text{ad}}$ the associated control $u_\mu^\star = \sum_{j=1}^{N(\varepsilon)} \alpha_\mu^j u_j$ yields

$$\|x_\mu^\star(t_\text{f}) - x^1\| \le \varepsilon \quad \text{for any chosen } \mu \in \mathcal{D}_{\text{ad}}$$

for the state $x = x_\mu^\star$ solving $\dot{x}(t) = \mathrm{A}_\mu x(t) + \mathrm{B}_\mu u_\mu^\star(t)$. $\rightarrow N(\varepsilon)$ a.s$_{mall}$ a.p.

Greedy controllability – Offline

For given tolerance $\varepsilon > 0$ select (with a greedy approach) $\mu_1, \ldots, \mu_{N(\varepsilon)} \in \mathcal{D}_{\text{ad}}^{\text{gr}}$ ($\mathcal{D}_{\text{ad}}^{\text{gr}} \subset \mathcal{D}_{\text{ad}}$ discrete training set) and compute associated $u_1, \ldots, u_{N(\varepsilon)}$.

Greedy controllability – Online

For any $\mu \in \mathcal{D}_{\text{ad}}$ set $u_\mu^\star = \sum_{j=1}^{N(\varepsilon)} \alpha_\mu^j u_j$ and solve $\dot{x}_\mu^\star(t) = \mathrm{A}_\mu x_\mu^\star(t) + \mathrm{B}_\mu u_\mu^\star(t) \Rightarrow \|x_\mu^\star(t_\text{f}) - x^1\| \le \varepsilon$

$\rightarrow \alpha_\mu^j$ projection of $x^1 - e^{t_\text{f}\mathrm{A}_\mu} x^\circ$ onto the j-th basis element for $\mathrm{span}\{\Lambda_\mu \varphi_1^\circ, \ldots, \Lambda_\mu \varphi_{N(\varepsilon)}^\circ\}$

Selection of the parameter values

Linear-quadratic optimal control problem:

$$\begin{cases} \min J_\mu(\varphi,\varphi^\circ) = \frac{1}{2}\int_0^{t_\mathsf{f}} \|\mathrm{B}_\mu^\top \varphi(t)\|^2\,\mathrm{d}t + \langle x^1,\varphi^\circ\rangle + \langle x^\circ,\varphi(0)\rangle \\ \text{s.t. } \dot\varphi(t) = -\mathrm{A}_\mu^\top\varphi(t) \text{ for } t\in[0,t_\mathsf{f}) \quad \text{and} \quad \varphi(t_\mathsf{f}) = \varphi^\circ \end{cases} \tag{$\mathbf{QP}_\mu$}$$

Control choice: $u_\mu = \mathrm{B}_\mu^\top\varphi_\mu$ steers the state $x = x_\mu$ to the desired target x^1

Greedy algorithm (offline)

[1] Choose $\mu_1 \in \mathcal{D}_{\mathsf{ad}}^{\mathsf{gr}}$;
[2] Compute $(\varphi_1,\varphi_1^\circ)$ by solving $(\mathbf{QP}_\mu)$ with $\mu=\mu_1$; set $u_1 = \mathrm{B}_{\mu_1}^\top\varphi_1$, $\Phi_1^\circ = \mathrm{span}\{\varphi_1^\circ\}$;
[3] Find $\mu_2 = \arg\max\{\mathrm{dist}(\varphi_\mu^\circ,\Phi_1^\circ)\,|\,\mu\in\mathcal{D}_{\mathsf{ad}}^{\mathsf{gr}}\}$;
[4] Compute $(\varphi_2,\varphi_2^\circ)$ by solving $(\mathbf{QP}_\mu)$ with $\mu=\mu_2$; set $u_2 = \mathrm{B}_{\mu_2}^\top\varphi_2$, $\Phi_2^\circ = \mathrm{span}\{\varphi_1^\circ,\varphi_2^\circ\}$;
⋮
[n] **until** $\mathrm{dist}(\varphi_\mu^\circ,\Phi_N^\circ) \le \epsilon$ for all $\mu\in\mathcal{D}_{\mathsf{ad}}^{\mathsf{gr}}$

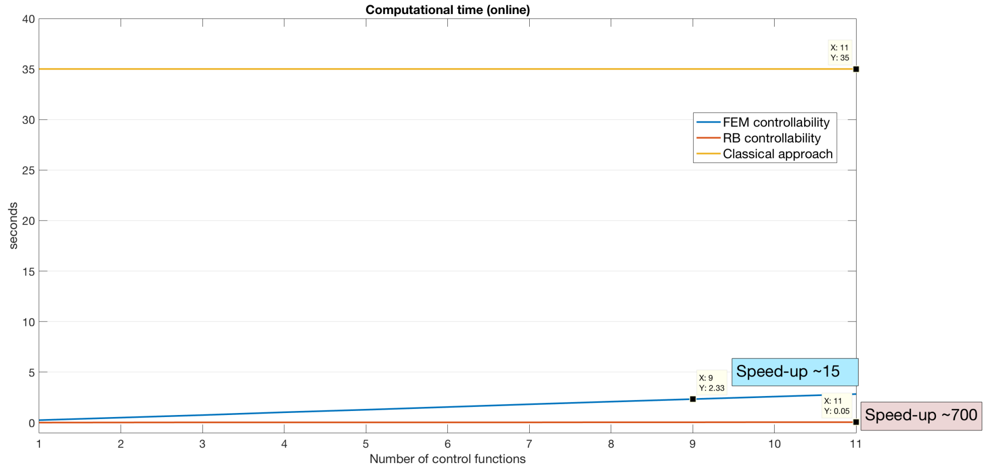
Numerical Example

State equation: for $\mu \in \mathcal{D}_{\text{ad}} = [0.5, 4] \subset \mathbb{R}$ consider

$$\begin{aligned} \dot{y}(t,\boldsymbol{x}) - \mu \Delta y(t,\boldsymbol{x}) + \beta \cdot \nabla y(t,\boldsymbol{x}) &= 0, & (t,\boldsymbol{x}) &\in Q = (0,t_{\text{f}}) \times \Omega \\ \mu \frac{\partial y}{\partial \boldsymbol{n}}(t,\boldsymbol{s}) &= u(t,\boldsymbol{s}), & (t,\boldsymbol{s}) &\in \Sigma = (0,t_{\text{f}}) \times \partial\Omega \\ y(0,\boldsymbol{x}) &= y^\circ(\boldsymbol{x}), & \boldsymbol{x} &\in \Omega = (0,2) \times (0,1) \subset \mathbb{R}^2 \end{aligned} \qquad (\mathbf{S}_\mu)$$

FE Galerkin discretization of $(\mathbf{S}_\mu)$: $\dot{x}(t) = \mathrm{A}_\mu x(t) + \mathrm{B}u(t)$ for $t \in (0, t_{\text{f}}]$, $\quad x(0) = x^\circ$

– utilize ROM for the dynamical systems
– weak greedy approximation property



5 Conclusions

ROM-based multiobjective optimization:

– multicriterial PDE-constrained optimization problems

– scalarization by the reference point method (for instance)

– apply (tailored and certified) reduced-order methods

– use of a-posteriori error analysis and basis update strategies to ensure required accuracy

Greedy controllability of parametrized linear systems:

– Greedy controllability approach for parametrized linear dynamical systems

– significant speed-up by the reduced-order greedy controllability

6 References

1: Banholzer'16: *POD-Based Bicriterial Optimal Control of Convection-Diffusion Equations*. Master thesis.
2: Banholzer/Beermann/V.'17: *POD-based error control for reduced-order bicriterial PDE-constrained optimization*.
3: Banholzer/V.'17: *Hierarchical Convex Multiobjective Optimization by the Euclidean Reference Point Method*.
4: Banholzer/Dellnitz/Gebken/Peitz/V.'19: *ROM-based multiobjective optimization of elliptic PDEs via numerical continuation*.
5: Beermann'19: *POD-Based Scalar and Multiobjective Optimal Control of Evolution Equations Including Operater-Valued Variables*. PhD thesis.
6: Beermann/Dellnitz/Peitz/V.'17: *POD-based multiobjective optimal control of PDEs with non-smooth objectives*.
7: Beermann/Dellnitz/Peitz/V.'18.: *Set-oriented multiobjective optimal control of PDEs using proper orthogonal decomposition*.
8: Banholzer/Makarov/Volkwein'18: *POD-based multiobjective optimal control of time-variant heat phenomena*.
9: Dellnitz/Schütze/Hestermeyer'05: *Covering pareto sets by multilevel subdivision techniques*.
10: Fabrini/Iapichino/V.'18: *Reduced-order greedy controllability of finite dimensional linear systems*
11: Gebken/Peitz/Dellnitz'18: *On the hierarchical structure of Pareto critical sets*.
12: Gebken/Peitz/Dellnitz'19: *A descent method for equality and inequality constrained multiobjective optimization problems*.
13: Iapichino/Trenz/V.'16: *Multiobjective optimal control of semilinear parabolic problems using POD*.
14: Iapichino/Ulbrich/V.'17: *Multiobjective PDE-constrained optimization using the reduced-basis method*.
15: Makarov'18: *POD-Based Bicriterial Optimal Control of Time-Dependent Convection-Diffusion Equations with Basis Update*. Master thesis.
16: Peitz'17: *Exploiting Structure in Multiobjective Optimization and Optimal Control*. PhD thesis.
17: Peitz/Dellnitz'18: *Gradient-based multiobjective optimization with uncertainties*.
18: Peitz/Ober-Blöbaum/Dellnitz'18: *Multiobjective optimal control methods for the Navier-Stokes equations using reduced order modeling*.
19: Spura'18: *Bicriterial Multiobjective Optimization Problems in 3D*. Master thesis.