

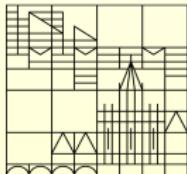
Surrogate Modeling in PDE Constrained Optimization

Stefan Volkwein

University of Konstanz, Department of Mathematics and Statistics, Numerics & Optimization Group

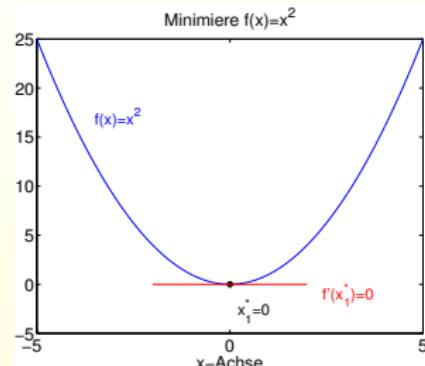
Projection Based Model Reduction:
RB Methods, POD, and Low Rank Tensor Approximations, November 23-29, 2014

Universität
Konstanz



Minimization with Inequality Constraints

- “Problem”:** $\min \{f(x) = x^2 : -\infty < x < \infty\}$
- Optimality condition:** $f'(x_1^*) \stackrel{!}{=} 0$ with $f'(x) = 2x$
- Solution:** $x_1^* = 0$



- Problem:** $\min \{f(x) = x^2 : 1 \leq x \leq 3\}$

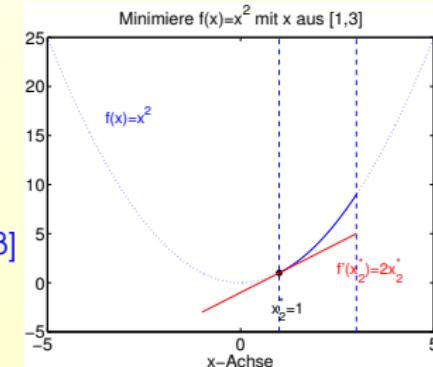
Lagrange functional:

$$L(x, \mu_a, \mu_b) = f(x) + \mu_a(1-x) + \mu_b(x-3)$$

Optimality condition:

$$\begin{pmatrix} f'(x_2^*) - \mu_a^* + \mu_b^* \\ \mu_a^*(1-x_2^*) \\ \mu_b^*(x_2^*-3) \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \underbrace{f'(x_2^*)(x - x_2^*)}_{2(x-1)} \geq 0 \quad \forall x \in [1, 3]$$

- Solution:** $x_2^* = 1, \mu_a^* = 2, \mu_b^* = 0$



Outline of Lecture 2

- Quadratic Programming (QP) Problems
- Extensions to Nonlinear PDE Constrained Optimization

Quadratic Programming (QP) Problems

Topics:

- Linear-quadratic optimal control problems
- POD-Galerkin schemes for first-order optimality system
- A-priori and a-posteriori error analysis
- Basis updates
- Regularized state constraints

Linear-Quadratic, Time-Variant Optimal Control Problem

- **Quadratic programming (QP) problem:**

$$\min_{x=(y,u)} J(x) = \frac{1}{2} \|y(t_f) - y_d\|_H^2 + \frac{\kappa}{2} \int_{t_o}^{t_f} \|u(t)\|_U^2 dt$$

subject to the linear evolution problem

$$\frac{d}{dt} \langle y(t), \varphi \rangle_H + a(t; y(t), \varphi) = \langle (f + \mathcal{B}u)(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f]$$

with $y(t_o) = y_o$ in H and to bilateral control constraints

$$u \in \mathcal{U}_{ad} = \{ \tilde{u} \in \mathcal{U} \mid u_a(t) \leq \tilde{u}(t) \leq u_b(t) \text{ a.e. in } [t_o, t_f] \}$$

- **State:** $y(t) \in V \hookrightarrow H$ with Hilbert spaces V, H
- **Control (Hilbert) space:** $\mathcal{U} = L^2(t_o, t_f; U)$ with $U = \mathbb{R}^{N_u}$, $U = L^2(\Omega)$ or $U = L^2(\Gamma)$
- **Input/control:** $u \in \mathcal{U}_{ad}$ (boundary or distributed control)
- **Bilinear form:** $a(t; \cdot, \cdot)$ continuous and $a(t; \varphi, \varphi) \geq \gamma_1 \|\varphi\|_V^2 - \gamma_2 \|\varphi\|_H^2$
- **Control operator:** $\mathcal{B} : \mathcal{U} \rightarrow L^2(t_o, t_f; V')$ linear, bounded
- Applicable also for **elliptic control problems** (Kahlbacher/V.'12, Tonn/Urban/V.'11, Tröltzsch/V.'09)

First-Order Necessary and Sufficient Optimality Conditions

- Quadratic programming (QP) problem:

$$\min_{x=(y,u)} J(x) = \frac{1}{2} \|y(t_f) - y_d\|_H^2 + \frac{\kappa}{2} \int_{t_o}^{t_f} \|u(t)\|_U^2 dt$$

$$\text{s.t. } \frac{d}{dt} \langle y(t), \varphi \rangle_H + a(t; y(t), \varphi) = \langle (f + \mathcal{B}u)(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f)$$

$$y(t_o) = y_o \text{ in } H \quad \text{and} \quad u_a(t) \leq u(t) \leq u_b(t), \quad t \in [t_o, t_f]$$

- Optimal state \bar{y} , optimal control $\bar{u} \in \mathcal{U}_{ad} = \{u \mid u_a \leq u \leq u_b \text{ in } [t_o, t_f]\}$

- Adjoint/dual equation:

$$-\frac{d}{dt} \langle \bar{p}(t), \varphi \rangle_H + a(t; \varphi, \bar{p}(t)) = 0 \quad \forall \varphi \in V \text{ a.e. in } [t_o, t_f], \quad \bar{p}(t_f) = \bar{y}(t_f) - y_d$$

- Variational inequality:

$$\int_{t_o}^{t_f} \langle \kappa \bar{u}(t) - (\mathcal{B}^* \bar{p})(t), u(t) - \bar{u}(t) \rangle_U dt \geq 0 \quad \forall u \in \mathcal{U}_{ad} \tag{VI}$$

- Reduced cost: $\hat{J}(u) = J(\hat{y} + \mathcal{S}u, u)$ with $\hat{J}'(\bar{u}) = \kappa \bar{u} - \mathcal{B}^* \bar{p} \in \mathcal{U}$, i.e.,

$$\langle \hat{J}'(\bar{u}), u(t) - \bar{u}(t) \rangle_{\mathcal{U}} \geq 0 \quad \forall u \in \mathcal{U}_{ad}$$

POD Galerkin for the State Variable

- **Particular solution:** $\hat{y} \in W(t_o, t_f)$ solves $\hat{y}(t_o) = y_o$ in H and

$$\frac{d}{dt} \langle \hat{y}(t), \varphi \rangle_H + a(t; \hat{y}(t), \varphi) = \langle f(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f]$$

- **POD space:** $V^\ell = \text{span}\{\psi_1, \dots, \psi_\ell\} \subset V$

- **POD control-to-state mapping:** $\mathcal{S}^\ell : \mathcal{U} \rightarrow W(t_o, t_f)$, $w^\ell = \mathcal{S}^\ell u$ solves $w^\ell(t_o) = 0$ in H and

$$\frac{d}{dt} \langle w^\ell(t), \psi \rangle_H + a(t; w^\ell(t), \psi) = \langle (\mathcal{B}u)(t), \psi \rangle_{V', V} \quad \forall \psi \in V^\ell \text{ a.e. in } (t_o, t_f]$$

$\Rightarrow \mathcal{S}^\ell$ linear and bounded

- **POD state:** $y^\ell = \hat{y} + \mathcal{S}^\ell u$

- **Previous theorem with a-priori results for the state:**

- $y^\ell(0) = y_o$ in H , i.e., no POD error in the initial condition
- $\|(\mathcal{S} - \mathcal{S}^\ell)u\|_{W(t_o, t_f)}^2 \leq \sum_{i>\ell} \lambda_i(u)$
- $\|\mathcal{S} - \mathcal{S}^\ell\|_{\mathcal{L}(\mathcal{U}, W(t_o, t_f))} \rightarrow 0$ for $\ell \rightarrow \infty$

POD Galerkin for the Dual Variable (Balancing POD)

- **Adjoint/dual equation:**

$$-\frac{d}{dt} \langle p(t), \varphi \rangle_H + a(t; \varphi, p(t)) = 0 \quad \forall \varphi \in V \text{ a.e. in } [t_o, t_f], \quad p(t_f) = y(t_f) - y_d$$

- **Terminal condition:** $p(t_f) = y(t_f) - y_d = (\hat{y} + \mathcal{S}u)(t_f) - y_d = \hat{y}(t_f) - y_d + (\mathcal{S}u)(t_f)$

- **Particular solution:** $\hat{p} \in W(t_o, t_f)$ solves

$$-\frac{d}{dt} \langle \hat{p}(t), \varphi \rangle_H + a(t; \varphi, \hat{p}(t)) = 0 \quad \forall \varphi \in V \text{ a.e. in } [t_o, t_f], \quad \hat{p}(t_f) = \hat{y}(t_f) - y_d$$

- **Dual solution operator:** $\mathcal{A} : \mathcal{U} \rightarrow W(t_o, t_f)$, $v = \mathcal{A}u$ solves

$$-\frac{d}{dt} \langle v(t), \varphi \rangle_H + a(t; \varphi, v(t)) = 0 \quad \forall \varphi \in V \text{ a.e. in } [t_o, t_f], \quad v(t_f) = (\mathcal{S}u)(t_f)$$

- **POD dual solution operator:** $\mathcal{A}^\ell : \mathcal{U} \rightarrow W(t_o, t_f)$, $v^\ell = \mathcal{A}^\ell u$ solves

$$-\frac{d}{dt} \langle v^\ell(t), \psi \rangle_H + a(t; \psi, v^\ell(t)) = 0 \quad \forall \psi \in V^\ell \text{ a.e. in } [t_o, t_f], \quad v^\ell(t_f) = (\mathcal{S}^\ell u)(t_f) \in V^\ell$$

⇒ same POD basis for state and adjoint variable

POD A-Priori Analysis for the Dual Variable

- Continuous variant of POD: for $\ell \in \{1, \dots, d\}$ solve

$$\min \sum_{k=1}^{\rho} \int_{t_0}^{t_f} \left\| y^k(t) - \sum_{i=1}^{\ell} \langle y^k(t), \psi_i \rangle_X \psi_i \right\|_X^2 dt \text{ s.t. } \{\psi_i\}_{i=1}^{\ell} \subset X, \langle \psi_i, \psi_j \rangle_X = \delta_{ij}, 1 \leq i, j \leq \ell \quad (\mathbf{P}_{\infty}^{\ell})$$

Theorem (Hinze/V.'08, Tröltzsch/V.'09, Gubisch/V.'13)

$$X = V, \mathcal{V} = \text{span}\{y^k(t) | t \in [t_0, t_f], 1 \leq k \leq \rho\}$$

- a) Snapshots $y^1 = \mathcal{S}u, y^2 = \mathcal{A}u, p = \hat{p} + \mathcal{A}u, p^{\ell} = \hat{p} + \mathcal{A}^{\ell}u$:

$$\|p - p^{\ell}\|_{W(t_0, t_f)}^2 \leq C_1 \left(\sum_{i>\ell} \lambda_i + \|y_t^1 - \mathcal{P}^{\ell} y_t^1\|_{L^2(t_0, t_f; V)}^2 + \|y_t^2 - \mathcal{P}^{\ell} y_t^2\|_{L^2(t_0, t_f; V')}^2 \right)$$

- b) Snapshots $y^1 = \mathcal{S}u, y^2 = \mathcal{A}u, y^3 = (\mathcal{S}u)_t, y^4 = (\mathcal{A}u)_t$, all in $L^2(t_0, t_f; V)$:

$$\|p - p^{\ell}\|_{W(t_0, t_f)}^2 \leq C_2 \sum_{i>\ell} \lambda_i$$

- c) If $\mathcal{S}\tilde{u}, \mathcal{A}\tilde{u} \in H^1(t_0, t_f; V)$ for all $\tilde{u} \in \mathcal{U}$, then $\lim_{\ell \rightarrow \infty} \|\mathcal{A} - \mathcal{A}^{\ell}\|_{\mathcal{L}(\mathcal{U}, W(t_0, t_f))} = 0$

In particular, $\lim_{\ell \rightarrow \infty} \|p(\tilde{u}) - p^{\ell}(\tilde{u})\|_{W(t_0, t_f)} = 0$ for any $\tilde{u} \in \mathcal{U}$

POD Approximation of the Variational Inequality

- **Variational inequality:**

$$\int_{t_0}^{t_f} \langle \kappa \bar{u}(t) - (\mathcal{B}^* \bar{p})(t), u(t) - \bar{u}(t) \rangle_U dt \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (\text{VI})$$

- **Optimal POD solutions:** $\bar{u}^\ell \in \mathcal{U}_{ad}$, $\bar{y}^\ell = \hat{y} + \mathcal{S}^\ell \bar{u}^\ell$, $\bar{p}^\ell = \hat{p} + \mathcal{A}^\ell \bar{u}^\ell$

- **POD variational inequality:**

$$\int_{t_0}^{t_f} \langle \kappa \bar{u}^\ell(t) - (\mathcal{B}^* \bar{p}^\ell)(t), u(t) - \bar{u}^\ell(t) \rangle_U dt \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (\text{VI}^\ell)$$

- **A-priori analysis:** choose $u = \bar{u}^\ell$ in (VI), $u = \bar{u}$ in (VI $^\ell$) and add

$$\begin{aligned} 0 &\leq \int_{t_0}^{t_f} \langle \kappa(\bar{u} - \bar{u}^\ell)(t) - (\mathcal{B}^*(\bar{p} - \bar{p}^\ell))(t), \bar{u}^\ell(t) - \bar{u}(t) \rangle_U dt \\ &= -\kappa \|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}}^2 - \underbrace{\int_{t_0}^{t_f} \langle (\mathcal{B}^*(\bar{p} - \bar{p}^\ell))(t), \bar{u}^\ell(t) - \bar{u}(t) \rangle_U dt}_{\leq C \|(\mathcal{A} - \mathcal{A}^\ell)\bar{u}\|_{L^2(t_0, t_f; V)} \|\bar{u}^\ell - \bar{u}\|_{\mathcal{U}}} \end{aligned}$$

Convergence Result for the POD Suboptimal Control

- Continuous variant of POD: for $\ell \in \{1, \dots, d\}$ solve

$$\min \sum_{k=1}^{\mathcal{P}} \int_{t_0}^{t_f} \left\| y^k(t) - \sum_{i=1}^{\ell} \langle y^k(t), \psi_i \rangle_X \psi_i \right\|_X^2 dt \text{ s.t. } \{\psi_i\}_{i=1}^{\ell} \subset X, \langle \psi_i, \psi_j \rangle_X = \delta_{ij}, 1 \leq i, j \leq \ell \quad (\mathbf{P}_\infty^\ell)$$

Theorem (Hinze/V.'08, Tröltzsch/V.'09, Gubisch/V.'13)

$$X = V, \mathcal{V} = \text{span}\{y^k(t) \mid t \in [t_0, t_f], 1 \leq k \leq \mathcal{P}\}$$

a) Snapshots $y^1 = \mathcal{S}\bar{u}, y^2 = \mathcal{A}\bar{u}, y^3 = (\mathcal{S}\bar{u})_t, y^4 = (\mathcal{A}\bar{u})_t$, all in $L^2(t_0, t_f; V)$:

$$\|\bar{u} - \bar{u}^\ell\|_{W(t_0, t_f)}^2 \leq C_2 \sum_{i>\ell} \lambda_i$$

b) If $\mathcal{S}\tilde{u}, \mathcal{A}\tilde{u} \in H^1(t_0, t_f; V)$ for all $\tilde{u} \in \mathcal{U}$, then $\lim_{\ell \rightarrow \infty} \|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}} = 0$

Perturbation Analysis (Malanowski/Büskens/Maurer'97)

● **Variational inequality:**

$$\int_{t_0}^{t_f} \langle \kappa \bar{u}(t) - (\mathcal{B}^* \bar{p})(t), u(t) - \bar{u}(t) \rangle_U dt \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (\text{VI})$$

● **Misfit in the variational inequality:** suboptimal $\bar{u}^\ell \in \mathcal{U}_{ad}$

$$\int_{t_0}^{t_f} \langle \kappa \bar{u}^\ell(t) - (\mathcal{B}^* \tilde{p}^\ell)(t), u(t) - \bar{u}^\ell(t) \rangle_U dt \not\geq 0 \quad \forall u \in \mathcal{U}_{ad}$$

with $\tilde{p}^\ell = \hat{p} + \mathcal{A} \bar{u}^\ell$

● **Perturbation analysis:** there exists a perturbation $\zeta^\ell \in \mathcal{U}$ satisfying

$$\int_{t_0}^{t_f} \langle \kappa \bar{u}^\ell(t) - (\mathcal{B}^* \tilde{p}^\ell)(t) + \zeta^\ell(t), u(t) - \bar{u}^\ell(t) \rangle_U dt \geq 0 \quad \forall u \in \mathcal{U}_{ad} \quad (\tilde{\text{VI}}^\ell)$$

● **A-posteriori analysis:** choose $u = \bar{u}^\ell$ in (VI), $u = \bar{u}$ in ($\tilde{\text{VI}}^\ell$) and add

$$\kappa \|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}}^2 \leq \int_{t_0}^{t_f} \langle (\mathcal{B}^* \mathcal{A}(\bar{u}^\ell - \bar{u})) + \zeta^\ell(t), \bar{u}^\ell(t) - \bar{u}(t) \rangle_U dt$$

since $\tilde{p}^\ell - \bar{p} = \mathcal{A}(\bar{u}^\ell - \bar{u})$

● **Estimate for the control:** $\|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}} \leq \frac{1}{\kappa} \|\zeta^\ell\|_{\mathcal{U}}$

Convergence Result for the Perturbation

- **Estimate for the control:** $\|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}} \leq \frac{1}{\kappa} \|\zeta^\ell\|_{\mathcal{U}}$ and $\tilde{p}^\ell = \hat{p} + \mathcal{A}\bar{u}^\ell$
 - **Computation of ζ^ℓ :** $\zeta^\ell(t) = \begin{cases} -(\kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } u_a(t) < \bar{u}^\ell(t) < u_b(t) \\ -\min(0, \kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } \bar{u}^\ell(t) = u_a(t) \\ -\max(0, \kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } \bar{u}^\ell(t) = u_b(t) \end{cases}$
- i.e., $\zeta^\ell = \zeta^\ell(\bar{u}^\ell) \Rightarrow \text{a-posteriori error analysis for suboptimal } \bar{u}^\ell$

Theorem (Tröltzsch/V.'09, Gubisch/V.'13)

$$X = V, \mathcal{V} = \text{span} \{y^k(t) \mid t \in [t_o, t_f], 1 \leq k \leq \mathcal{P}\}$$

a) Snapshots $y^1 = \mathcal{S}\bar{u}, y^2 = \mathcal{A}\bar{u}, y^3 = (\mathcal{S}\bar{u})_t, y^4 = (\mathcal{A}\bar{u})_t$, all in $L^2(t_o, t_f; V)$:

$$\|\zeta^\ell\|_{\mathcal{U}}^2 \leq C \sum_{i>\ell} \lambda_i$$

b) If $\mathcal{S}\tilde{u}, \mathcal{A}\tilde{u} \in H^1(t_o, t_f; V)$ for all $\tilde{u} \in \mathcal{U}$, then $\lim_{\ell \rightarrow \infty} \|\zeta^\ell\|_{\mathcal{U}} = 0$

Algorithm with POD A-Posteriori Analysis

- **Estimate for the control:** $\|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}} \leq \frac{1}{\kappa} \|\zeta^\ell\|_{\mathcal{U}}$ and $\tilde{p}^\ell = \hat{p} + \mathcal{A}\bar{u}^\ell$
- **Computation of ζ^ℓ :** $\zeta^\ell(t) = \begin{cases} -(\kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } u_a(t) < \bar{u}^\ell(t) < u_b(t) \\ -\min(0, \kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } \bar{u}^\ell(t) = u_a(t) \\ -\max(0, \kappa\bar{u}^\ell(t) - (\mathcal{B}^*\tilde{p}^\ell)(t)) & \text{if } \bar{u}^\ell(t) = u_b(t) \end{cases}$

Algorithmus 1 (Optimal control with a-posteriori error estimation)

```
(1) Choose  $\ell_{\max}$  and POD basis  $\{\psi_i\}_{i=1}^{\ell}$  for the Galerkin approximation of the LQ problem;
(2) Determine the reduced-order model for the LQ problem;
(3) Calculate suboptimal control  $\bar{u}^\ell \in \mathcal{U}_{ad}$ , e.g., by a semismooth Newton method;
(4) Compute perturbation  $\bar{\zeta}^\ell = \zeta^\ell(\bar{u}^\ell)$ ;
(5) IF  $\|\bar{\zeta}^\ell\|_{\mathcal{U}}/\kappa > \text{TOL}$  AND  $\ell < \ell_{\max}$  THEN
    Enlarge  $\ell$  and go back to (2);
ELSE
    Stop;
ENDIF
```

- Applicable for **balanced-truncation** (Vossen/V.'12) Or **reduced-basis method** (Tonn/Urban/V.'11)
- Error estimation between **high- and low-dimensional discretization** (Gubisch/Neitzel)

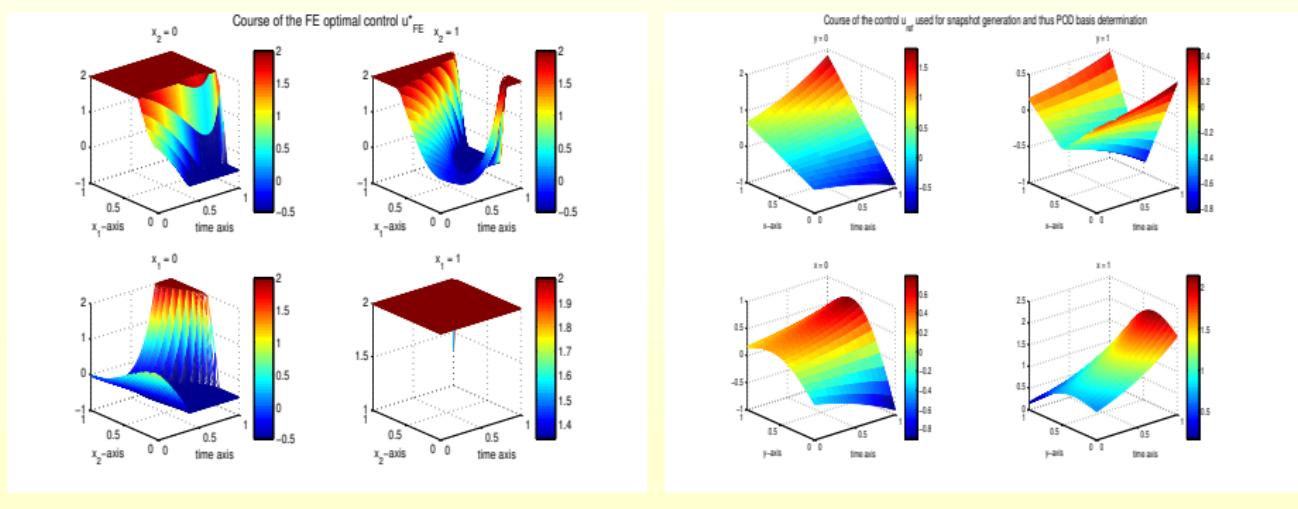
Numerical Example: Problem Formulation and Optimal Control (Studerger'12, Studerger/V.'13)

Consider:

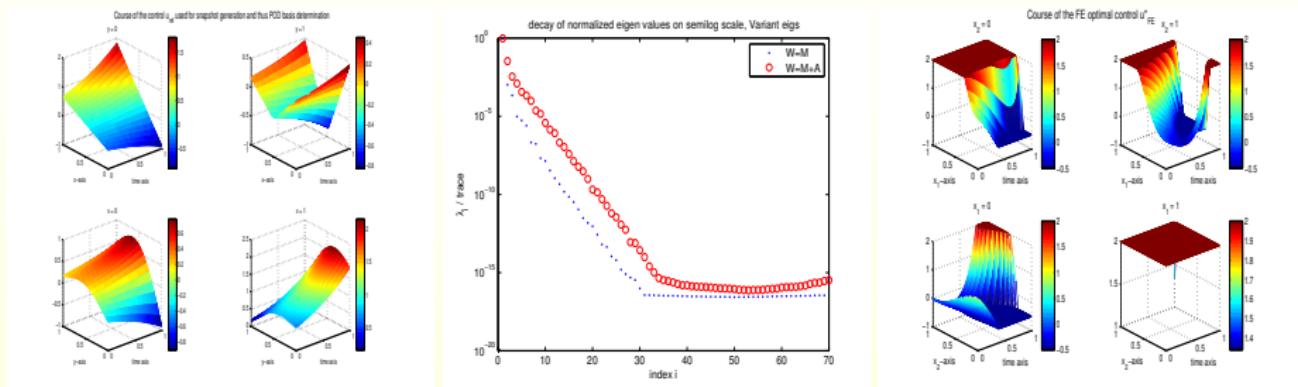
$$\min J(y, u) = \frac{1}{2} \int_{\Omega} |y(t_f) - y^d|^2 dx + \frac{1}{200} \int_0^{t_f} \int_{\Gamma} |u|^2 d\Gamma dt$$

$$\text{s.t. } y_t + 0.1\Delta y = 0 \text{ in } Q, \quad \frac{\partial y}{\partial n} + \frac{y}{100} = u \text{ on } \Sigma, \quad y(0) = y_0 \text{ in } \Omega = (0, 1)^2 \\ -0.5 \leq u \leq 2 \text{ on } \Sigma = (0, t_f) \times \Gamma$$

Method & Discretization: semismooth Newton & implizit Euler, finite elements



Numerical Example: POD Error Analysis (Studinger'12, Studinger/V.'13)



A-posteriori error: $\|\bar{u} - \bar{u}^\ell\|_{\mathcal{U}} \leq \frac{1}{\kappa} \|\zeta^\ell(\bar{u}^\ell)\|_{\mathcal{U}} =: \varepsilon_{\text{ape}}$

ℓ	ε_{ape}	$\ \bar{u}^h - \bar{u}^\ell\ $	$\frac{\varepsilon_{\text{ape}}}{\ \bar{u} - \bar{u}^\ell\ }$	ε_{ape}	$\ \bar{u}^h - \bar{u}^\ell\ $	$\frac{\varepsilon_{\text{ape}}}{\ \bar{u} - \bar{u}^\ell\ }$
5	$1.3 \cdot 10^{-0}$	$9.1 \cdot 10^{-1}$	1.32	$6.5 \cdot 10^{-1}$	$5.6 \cdot 10^{-1}$	1.16
20	$5.9 \cdot 10^{-1}$	$3.2 \cdot 10^{-1}$	1.84	$7.5 \cdot 10^{-3}$	$7.3 \cdot 10^{-3}$	1.03
60	$1.4 \cdot 10^{-2}$	$1.2 \cdot 10^{-2}$	1.17	$8.3 \cdot 10^{-5}$	$8.3 \cdot 10^{-5}$	1.00
70	$1.2 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$	1.10	$3.0 \cdot 10^{-5}$	$3.0 \cdot 10^{-5}$	1.00
90	$1.1 \cdot 10^{-2}$	$9.7 \cdot 10^{-3}$	1.13	$3.7 \cdot 10^{-6}$	$3.7 \cdot 10^{-6}$	1.00

Problem: optimal FE control \bar{u}^h is unknown \Rightarrow **POD basis update**

Optimality-System POD (OS-POD) (Kunisch/V.'08)

- Original problem: $\min \hat{J}(u) = J(y(u), u)$ s.t. $u \in \mathcal{U}_{ad}$

- POD-Galerkin approximation:

$$\min \hat{J}^\ell(u) = J(y^\ell(u), u) \quad \text{s.t.} \quad u \in \mathcal{U}_{ad} \quad (\hat{\mathbf{P}}^\ell)$$

- OSPOD-Problem:

$$\min \hat{\mathcal{J}}^\ell(u) = J(y^\ell(u), u) \quad \text{s.t.} \quad \begin{cases} y = y(u), \quad \psi_i = \psi_i(u) \text{ for } 1 \leq i \leq \ell \\ \mathcal{R}\psi_i = \int_{t_0}^{t_f} \langle y(t), \psi_i \rangle_X y(t) dt = \lambda_i \psi_i, \quad 1 \leq i \leq \ell \end{cases} \quad (\hat{\mathbf{P}}_{\text{ospod}}^\ell)$$

→ more complex than the original problem

- Numerical realization: operator splitting

- choose an initial control $u^{(0)}$ and a corresponding POD basis $\{\psi_i^{(0)}\}_{i=1}^\ell$
- improve the POD basis by applying a few gradient projection steps for $(\hat{\mathbf{P}}_{\text{ospod}}^\ell)$
- compute an approximate solution to $(\hat{\mathbf{P}}^\ell)$ by Algorithm 1

- Efficient combination with a-posteriori error estimator (Grimm'13, Grimm/Gubisch/V.'14, V.'11)

Algorithm for A-Posteriori Error with OS-POD (Grimm'13, Grimm/Gubisch/V.'14, V.'11)

- OS-POD-Problem:

$$\min \hat{\mathcal{J}}^\ell(u) = J(y^\ell(u), u) \quad \text{s.t.} \quad \begin{cases} y = y(u), \quad \psi_i = \psi_i(u) \text{ for } 1 \leq i \leq \ell \\ \mathcal{R}\psi_i = \int_{t_0}^{t_f} \langle y(t), \psi_i \rangle_X y(t) dt = \lambda_i \psi_i, \quad 1 \leq i \leq \ell \end{cases} \quad (\hat{\mathbf{P}}_{\text{ospod}}^\ell)$$

Algorithmus 1 (*Optimal control with a-posteriori error estimation*)

- (1) Choose initial control u^0 and number of basis elements $\ell < \ell_{\max}$;
- (2) Perform k OS-POD gradient steps to get u^k and $\{\psi_i(u^k)\}_{i=1}^\ell$;

- (3) Determine the reduced-order model for the LQ problem;
- (4) Calculate suboptimal control $\bar{u}^\ell \in \mathcal{U}_{ad}$, e.g., by a semismooth Newton method;
- (5) Compute perturbation $\bar{\zeta}^\ell = \zeta^\ell(\bar{u}^\ell)$;
- (6) **IF** $\|\bar{\zeta}^\ell\|_{\mathcal{U}} / \kappa > \text{TOL}$ **AND** $\ell < \ell_{\max}$ **THEN**
 Enlarge ℓ and go back to (3);
ELSE
 Stop;
ENDIF

- Algorithm:** apply a few gradient steps for $(\hat{\mathbf{P}}_{\text{ospod}}^\ell)$ and continue with fixed POD basis
- Stopping criterium:** $\|\bar{\zeta}^\ell\|_{\mathcal{U}} / \kappa = \mathcal{O}(\Delta t^\alpha + \Delta x^\beta)$

OS-POD for linear-quadratic, control constrained control (Grimm'13, Grimm/Gubisch/V'14, V'11)

Consider:

$$\min J(y, u) = \frac{1}{2} \int_{\Omega} |y(t_f) - y^d|^2 dx + \frac{1}{20} \int_0^{t_f} \int_{\Gamma} |u|^2 d\mathbf{x} dt$$

s.t. $y_t + 0.1\Delta y = 0$ in Q , $\frac{\partial y}{\partial n} + \frac{y}{100} = u$ on Σ , $y(0) = y_0$ in $\Omega = (0, 1)^2$
 $-0 \leq u \leq 1$ on $\Sigma = (0, t_f) \times \Gamma$

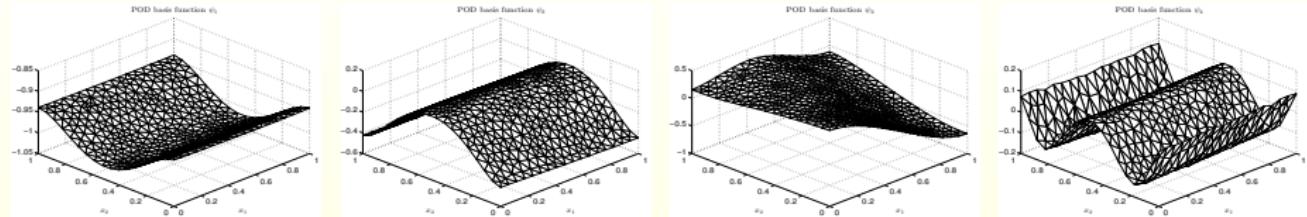
Method & Discretization: semismooth Newton & implizit Euler, finite elements

	$k = 0$	$k = 1$	$k = 2$	with u_{FE}
required ℓ	35	40	13	13
CPU	110.77 sec	147.14 sec	18.39 sec	11.48 sec
$\ \zeta^\ell\ /\sigma_u$	$3.43 \cdot 10^{-3}$	$1.14 \cdot 10^{-2}$	$2.82 \cdot 10^{-3}$	$1.94 \cdot 10^{-3}$
e_{abs}^u	$3.15 \cdot 10^{-3}$	$9.53 \cdot 10^{-3}$	$2.62 \cdot 10^{-3}$	$1.93 \cdot 10^{-3}$
e_{rel}^u	$2.45 \cdot 10^{-3}$	$7.73 \cdot 10^{-3}$	$2.15 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$
e_{abs}^y	$1.59 \cdot 10^{-2}$	$3.10 \cdot 10^{-3}$	$7.55 \cdot 10^{-4}$	$1.92 \cdot 10^{-4}$
e_{rel}^y	$5.75 \cdot 10^{-3}$	$1.12 \cdot 10^{-3}$	$2.73 \cdot 10^{-4}$	$6.97 \cdot 10^{-5}$

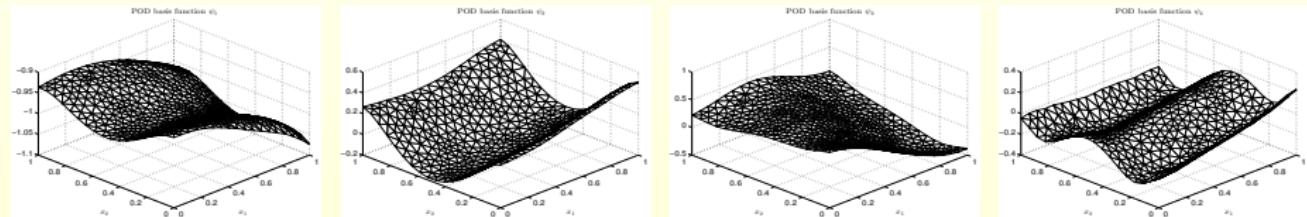
	$k = 0$	$k = 1$	$k = 2$	with u_{FE}
$\ \zeta^{15}\ /\sigma_u$	$2.50 \cdot 10^{-2}$	$1.45 \cdot 10^{-2}$	$2.27 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$
e_{abs}^u	$2.06 \cdot 10^{-2}$	$1.19 \cdot 10^{-2}$	$2.07 \cdot 10^{-3}$	$1.59 \cdot 10^{-3}$
e_{rel}^u	$1.65 \cdot 10^{-2}$	$9.65 \cdot 10^{-3}$	$1.67 \cdot 10^{-3}$	$1.30 \cdot 10^{-3}$
e_{abs}^y	$2.59 \cdot 10^{-2}$	$6.34 \cdot 10^{-3}$	$6.42 \cdot 10^{-4}$	$6.91 \cdot 10^{-5}$
e_{rel}^y	$9.34 \cdot 10^{-3}$	$2.29 \cdot 10^{-3}$	$2.32 \cdot 10^{-4}$	$2.51 \cdot 10^{-5}$
different u_a	96	67	15	16(2233)
different u_b	63	38	6	4 (3891)

Convergence of the POD basis (Grimm'13, Grimm/Gubisch/V.'14, Kunisch/Müller'14)

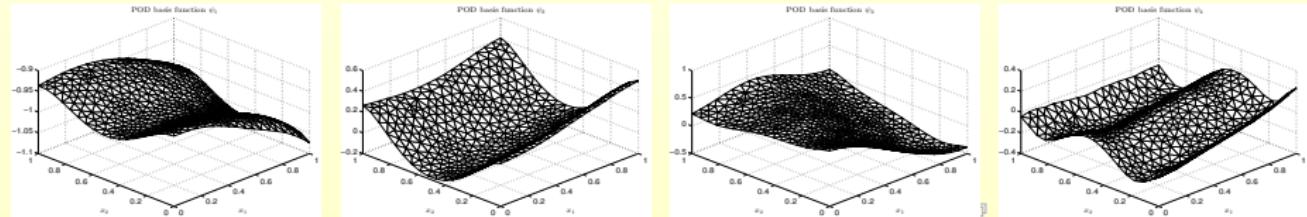
- First four POD basis functions generated from $u = 0$:



- First four POD basis functions generated from the optimal control \bar{u}^h :



- First four POD basis functions generated after $k = 2$ OS-POD gradient steps:



Problem with Mixed Control-State Constraints (Tröltzsch '05)

• (Abstract) linear-quadratic problem:

$$\min_{x=(y,u)} J(x) = \frac{1}{2} \|y(t_f) - y_d\|_H^2 + \frac{\kappa}{2} \int_{t_o}^{t_f} \|u(t)\|_U^2 dt$$

$$\text{s.t. } \frac{d}{dt} \langle y(t), \varphi \rangle_H + a(t; y(t), \varphi) = \langle (f + \mathcal{B}u)(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f]$$

$$y(t_o) = y_o \text{ in } H \quad \text{and} \quad u_a(t) \leq \varepsilon u(t) + (\mathcal{I}y)(t) \leq u_b(t), \quad t \in [t_o, t_f]$$

- Control space: $\mathcal{U} = L^2(t_o, t_f; U)$ with $U = \mathbb{R}^{N_u}$

- Input/control operator: $(\mathcal{B}u)(t, x) = \sum_{i=1}^{N_u} u_i(t) \chi_{\Omega_i}(x)$

- State operator: $(\mathcal{I}y)(t) = (\int_{\Omega_i} y(t, x) dx / |\Omega_i|)_{1 \leq i \leq N_u}$

- State constraints: $\varepsilon \rightarrow 0$

- Particular solution: $\hat{y} \in W(t_o, t_f)$ solves $\hat{y}(t_o) = y_o$ in H and

$$\frac{d}{dt} \langle \hat{y}(t), \varphi \rangle_H + a(t; \hat{y}(t), \varphi) = \langle f(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f]$$

- Control-to-state mapping: $\mathcal{S} : \mathcal{U} \rightarrow W(t_o, t_f)$, $w = \mathcal{S}u$ solves $w(t_o) = 0$ in H and

$$\frac{d}{dt} \langle w(t), \varphi \rangle_H + a(t; w(t), \varphi) = \langle (\mathcal{B}u)(t), \varphi \rangle_{V', V} \quad \forall \varphi \in V \text{ a.e. in } (t_o, t_f]$$

- Inequality constraints: $v_a(t) := (u_a - \mathcal{I}\hat{y})(t) \leq (\varepsilon u + \mathcal{I}\mathcal{S}u)(t) \leq (u_b - \mathcal{I}\hat{y})(t) =: v_b(t)$

Formulation as a Control Constrained Problem

- **Solution representation:** $u \mapsto y(u) = \hat{y} + \mathcal{S}u$
- **Inequality constraints:** $v_a(t) := (u_a - \mathcal{I}\hat{y})(t) \leq (\varepsilon u + \mathcal{I}\mathcal{S}u)(t) \leq (u_b - \mathcal{I}\hat{y}) =: v_b(t)$
- **Transformation of variables:** $v := \mathcal{F}u$ with $\mathcal{F} = \varepsilon + \mathcal{I}\mathcal{S}$, i.e. $y = \hat{y} + \mathcal{S}u = \hat{y} + \mathcal{S}\mathcal{F}^{-1}v$
- **Transformed, linear-quadratic problem:**

$$\begin{aligned} \min_v J(\hat{y} + \mathcal{S}\mathcal{F}^{-1}v, \mathcal{F}^{-1}v) &= \dots + \frac{\kappa}{2} \|\mathcal{F}^{-1}v\|^2 \\ \text{s.t. } v \in V_{ad} &= \{\tilde{v} \mid v_a(t) \leq \tilde{v}(t) \leq v_b(t) \text{ for } t \in [t_0, t_f]\} \end{aligned}$$

⇒ form of the previous linear-quadratic optimal control problem

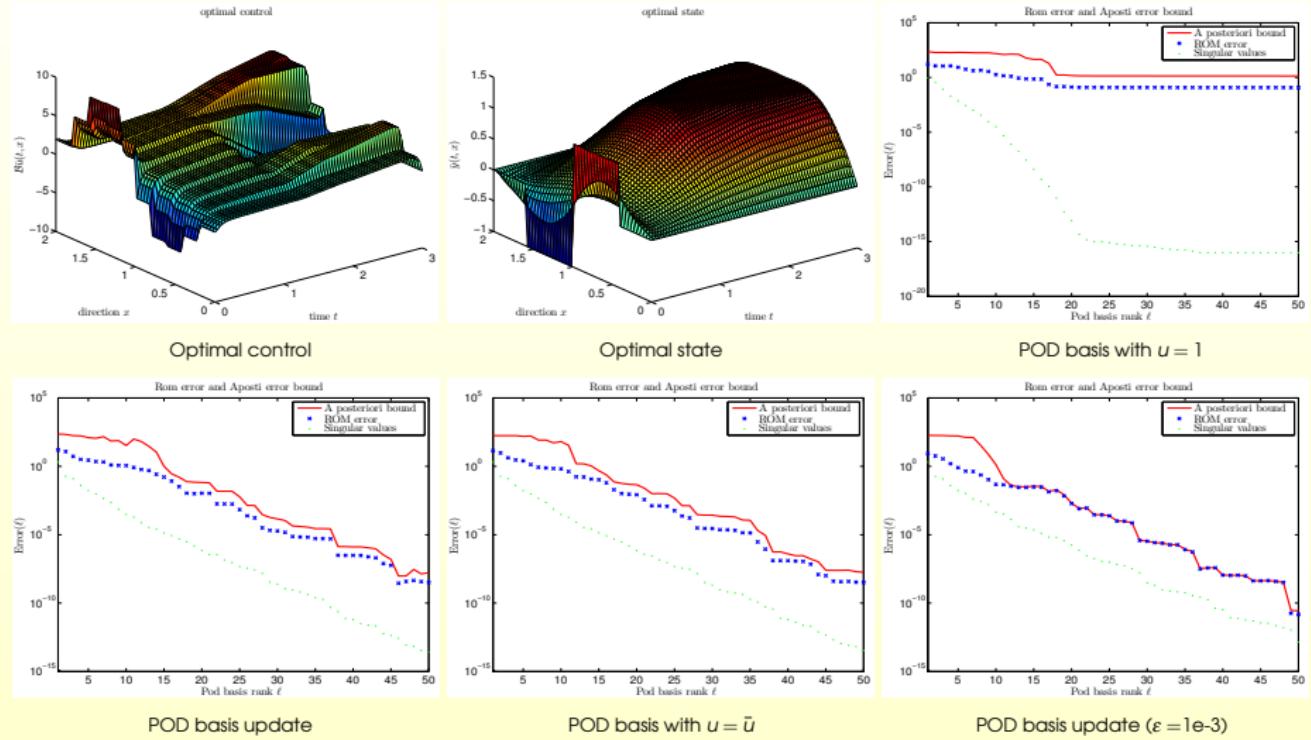
- **Variational inequality:**

$$\left\langle \kappa \mathcal{F}^{-\star} \mathcal{F}^{-1} \bar{v}(t) - \mathcal{B}^* \bar{p}(t), v(t) - \bar{v}(t) \right\rangle_{\mathcal{U}} \geq 0 \quad \forall v \in V_{ad}$$

- **POD Galerkin scheme:** additional analysis for $\mathcal{F}^\ell \approx \mathcal{F}$
- **A-posteriori error estimate** (Gubisch/V'14): $\|\bar{u} - \bar{u}^\ell\| \leq \frac{1}{\kappa} \|\mathcal{F}\|_{\mathcal{L}(\mathcal{U})} \|\zeta^\ell\|$ with computable ζ^ℓ
- **Convergence result** (Gubisch/V'14): $\|\zeta^\ell\| \rightarrow 0$ for $\ell \rightarrow \infty$

Numerical Example: POD basis update (Afanasiev/Hinze'01, Gubisch/V.'14)

Optimal control problem: heat equation, control space $\mathcal{U} = L^2(0,3; \mathbb{R}^{10})$, $\varepsilon = 1e-5$



State and Control Constrained Linear-Quadratic Optimal Control Problem

- Quadratic programming (QP) problem:

$$\min_{x=(y,u)} J(x) = \frac{1}{2} \|y(t_f) - y_d\|_{L^2(\Omega)}^2 + \frac{\kappa}{2} \int_{t_0}^{t_f} \|u(t)\|_U^2 dt$$

subject to the linear evolution problem

$$\langle y_t(t), \varphi \rangle + a(t; y(t))(t), \varphi \rangle = \langle (f + \mathcal{B}u)(t), \varphi \rangle \quad \forall \varphi \in V \text{ in } (t_0, t_f]$$

with $y(t_0) = y_0$ and to bilateral control constraints

$$\begin{aligned} u \in \mathcal{U}_{ad} &= \{v \in \mathcal{U} \mid u_a(t) \leq v(t) \leq u_b(t) \text{ in } [t_0, t_f]\} \\ y \in &\{z \in L^2(Q) \mid y_a(t, x) \leq z(t, x) \leq y_b(t, x) \text{ in } Q\} \end{aligned}$$

- Lavrentiev regularization: $\varepsilon > 0$ and $x = (y, u)$

$$J(x, w) = \frac{1}{2} \|y(t_f) - y_d\|_{L^2(\Omega)}^2 + \frac{\kappa}{2} \int_{t_0}^{t_f} \|u(t)\|_U^2 dt + \frac{\sigma}{2} \int_{t_0}^{t_f} \|w(t)\|_{L^2(\Omega)}^2 dt$$

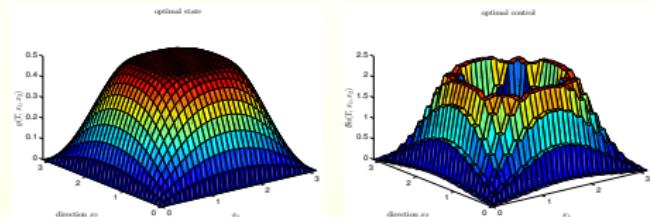
$$(y, w) \in \{(z, v) \in L^2(Q) \times L^2(Q) \mid y_a(t, x) \leq \varepsilon v(t, x) + z(t, x) \leq y_b(t, x) \text{ in } Q\}$$

- Regular Lagrange multipliers (Tröltzsch '05): formulation as a control constrained problem

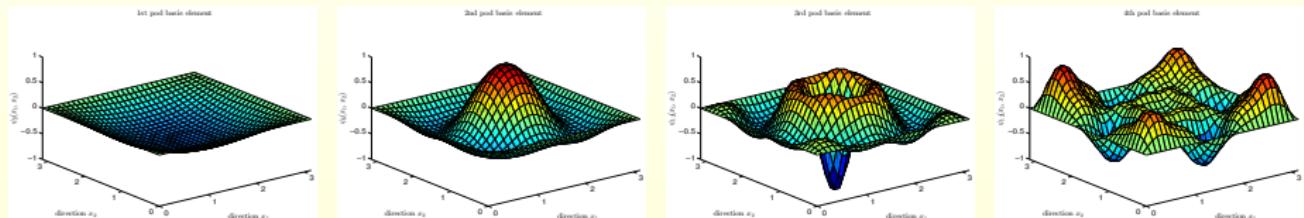
- A-posteriori error (Gubisch '14): extension of the analysis from the control-constrained case \Rightarrow OS-POD also applicable

Numerical Example: State and Control Constraints (Gubisch'14, Grimm/Gubisch/V.'14)

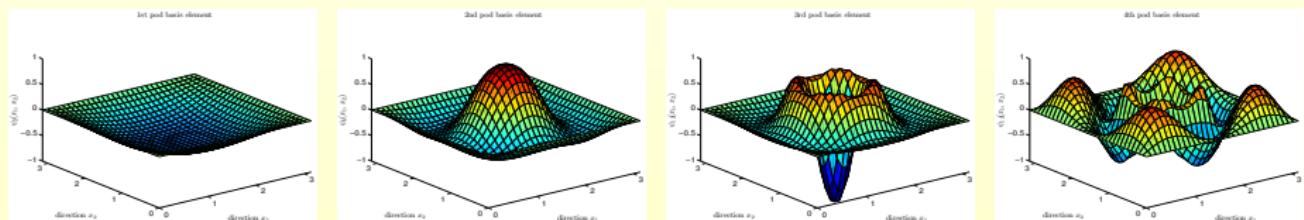
- Two-dimensional heat equation
- Distributed control
- Semismooth Newton method



- First four POD basis functions generated from the optimal control \bar{u} :



- First four POD basis functions generated after three OS-POD gradient steps:



Related Literature

- Afanasiev, Dede, Fahl, Grepl, Hinze, Kärcher, Manzoni, Negri, Quarteroni, Rozza, Sachs, Schu...
- E. Grimm: Optimality system POD and a-posteriori error analysis for linear-quadratic optimal control problems, 2013
- E. Grimm, M. Gubisch, S. V.: Numerical analysis of optimality-system POD for constrained optimal control, 2014
- M. Gubisch, S. V.: *POD a-posteriori error analysis for optimal control problems with mixed control-state constraints*, 2014
- M. Hinze, S. V.: *Error estimates for abstract linear-quadratic optimal control problems using POD*, 2008
- K. Kunisch, S. V.: *POD for optimality systems*, 2010
- A. Studinger, S.V.: *Numerical analysis of POD a-posteriori error estimation for optimal control*, 2013
- T. Tonn, K. Urban, S.V.: *Comparison of the reduced-basis and POD a-posteriori error estimators for an elliptic linear-quadratic optimal control problem*, 2011
- F. Tröltzsch, S. V.: *POD a-posteriori error estimates for linear-quadratic optimal control problems*, 2009
- S.V.: *Optimality system POD and a-posteriori error analysis for linear-quadratic problems*, 2011

Extensions to Nonlinear PDE Constrained Optimization

Topics:

- Inexact sequential quadratic programming (SQP)
- A-posteriori error analysis
- Trust-region POD

Sequential Quadratic Programming (SQP)

- **Infinite dimensional optimization:**

$$\min J(x) \quad \text{s.t.} \quad e(x) = 0, \quad g(x) \leq 0 \quad (\mathbf{P})$$

- **Lagrange functional for (P):** $\mathcal{L}(x, p, q) = J(x) + \langle e(x), p \rangle + \langle g(x), q \rangle$

- **(Local) SQP method:** at $z_k = (x_k, p_k, q_k)$ solve

$$\begin{cases} \min_{x_\delta} \mathcal{L}_x(z_k) x_\delta + \frac{1}{2} \mathcal{L}_{xx}(z_k)(x_\delta, x_\delta) \\ \text{s.t. } e(x_k) + e'(x_k)x_\delta = 0, \quad g(x_k) + g'(x_k)x_\delta \leq 0 \end{cases} \quad (\mathbf{QP}^k)$$

- **KKT system for equality constraints:** solution \bar{x}_δ to (\mathbf{QP}^k) is characterized by

$$\underbrace{\begin{pmatrix} \mathcal{L}_{xx}(z_k) & e'(x_k)^* \\ e'(x_k) & 0 \end{pmatrix}}_{A_k} \underbrace{\begin{pmatrix} \bar{x}_\delta \\ \bar{p}_\delta \end{pmatrix}}_{\bar{z}_\delta} = -\underbrace{\begin{pmatrix} \mathcal{L}_x(z_k) \\ e(x_k) \end{pmatrix}}_{b_k}$$

Inexact SQP by Using POD or RB

- **KKT system:** solution \bar{x}_δ to (\mathbf{QP}^k) is characterized by

$$\underbrace{\begin{pmatrix} \mathcal{L}_{xx}(z_k) & e'(x_k)^* \\ e'(x_k) & 0 \end{pmatrix}}_{A_k} \underbrace{\begin{pmatrix} \bar{x}_\delta \\ \bar{p}_\delta \end{pmatrix}}_{\bar{z}_\delta} = -\underbrace{\begin{pmatrix} \mathcal{L}_x(z_k) \\ e(x_k) \end{pmatrix}}_{b_k}$$

- **KKT system:** inexact solve of $A_k \bar{z}_\delta = b_k$ by discretization

- **Discretization:** (POD or RB or BT or...) model reduction

$$A_k^\ell \bar{z}_\delta^\ell = b_k^\ell \in \mathbb{R}^n, \quad n = n(\ell)$$

- **Convergence of (local) SQP method:** \bar{z}_δ^ℓ reduced-order solution

$$\|A_k \mathcal{P} \bar{z}_\delta^\ell - b_k\| = \mathcal{O}(\|\mathcal{L}'(z_k)\|^q), \quad q \in [1, 2]$$

with prolongation \mathcal{P}

- **Rate of convergence:** superlinear ($1 < q < 2$), quadratic ($q = 2$)

- **Control of reduced-order approach:**

$$\|A_k \mathcal{P} \bar{z}_\delta^\ell - b_k\| \simeq \|\bar{z}_\delta - \mathcal{P} \bar{z}_\delta^\ell\| \simeq \|\mathcal{L}'(z_k)\|^q$$

Convergence Result

- Variables in optimal control: $x = (y, u)$, $y = y(u)$
- KKT system: $z_k = (x_k, p_k)$, $x_k = (y_k, u_k)$

$$\left(\begin{array}{cc|c} \mathcal{L}_{yy}(z_k) & \mathcal{L}_{yu}(z_k) & e_y(x_k)^* \\ \mathcal{L}_{uy}(z_k) & \mathcal{L}_{uu}(z_k) & e_u(x_k)^* \\ \hline e_y(x_k) & e_u(x_k) & 0 \end{array} \right) \left(\begin{array}{c} y_\delta \\ u_\delta \\ p_\delta \end{array} \right) = \left(\begin{array}{c} -\mathcal{L}_y(z_k) \\ -\mathcal{L}_u(z_k) \\ -e(x_k) \end{array} \right)$$

- Suboptimal solution to KKT system: $\bar{z}_\delta^\ell = (\bar{y}_\delta^\ell, \bar{u}_\delta^\ell, \bar{p}_\delta^\ell)$
- Prolongation \mathcal{P} : $\bar{z}_\delta^\ell \mapsto \mathcal{P}\bar{z}_\delta^\ell = (\tilde{y}_\delta, \tilde{u}_\delta^\ell, \tilde{p}_\delta)$ with

$$e_y(x_k)\tilde{y}_\delta = -e(x_k) - e_u(x_k)\bar{u}_\delta^\ell$$

$$e_y(x_k)^*\tilde{p}_\delta = -\mathcal{L}_y(z_k) - \mathcal{L}_{yy}(z_k)\tilde{y}_\delta - \mathcal{L}_{yu}(z_k)\bar{u}_\delta^\ell$$

→ a-posteriori error computable

Theorem (Kahlbacher/V.'12)

Second-order sufficient optimality implies

$$\lim_{k \rightarrow \infty} z_k + \mathcal{P}\bar{z}_\delta^\ell = \bar{z} \quad \text{if} \quad \|A_k \mathcal{P}\bar{z}_\delta^\ell - b_k\| \simeq \|\bar{u}_\delta - \bar{u}_\delta^\ell\| < \text{TOL}$$

Multilevel Approach with Reduced-Order Models

- Convergence criterium: $\|A_k \mathcal{P} \bar{z}_\delta^\ell - b_k\| \simeq \|\bar{u}_\delta - \bar{u}_\delta^\ell\| < \text{TOL}$

- A-posteriori error (Tröltzsch/V.'09):

$$\|\bar{u}_\delta - \bar{u}_\delta^\ell\| \simeq \underbrace{\|\mathcal{L}_{uy}(z_k)\tilde{y}_\delta + \mathcal{L}_{uu}(z_k)\bar{u}_\delta^\ell + e_u(x_k)^* \tilde{p}_\delta + \mathcal{L}_u(z_k)\|}_{:= -\bar{\zeta}^\ell}$$

with $\|\bar{\zeta}^\ell\| \rightarrow 0$ for $\ell \rightarrow \infty$

- Convergence of $\|\bar{\zeta}^\ell\|$: no rate, basis dependent (Hinze/V.'08)

- POD basis: combination with Optimality-System POD (Metzdorf'15)

- Combination with FE adaptivity: (Clever/Lang/Ulbrich/Ziems)

- Reduced approach: $\min \hat{J}(u) = J(y(u), u)$

- QP problem: for given iterate u^k solve

$$\min q^k(u_\delta) = \hat{J}(u^k) + \hat{J}'(u^k)u_\delta + \frac{1}{2}\hat{J}''(u^k)(u_\delta, u_\delta)$$

- Globalization by inexact trust-region: inexact cost $q_\ell^k \approx q^k$ and inexact solution of

$$\min q_\ell^k(u_\delta) = \hat{J}_\ell(u^k) + \hat{J}'_\ell(u^k)u_\delta + \frac{1}{2}\hat{J}''_\ell(u^k)(u_\delta, u_\delta) \quad \text{s.t.} \quad \|u_\delta\| \leq \Delta_k$$

- Convergence criterium: Carter condition $\|\hat{J}'(u^k) - \hat{J}'_\ell(u^k)\| \leq \zeta \|\hat{J}'_\ell(u^k)\|$ with $\zeta \in (0, 1)$

- Trust-region POD: (Arian/Fahl/Sachs'00, Schu/Sachs'12) and (Rogg'14)

Related Literature

- Alla, Dihlmann, Haasdonk, Heinkenschloss, Hinze, Manzoni, Quarteroni, Rozza, Sachs, Ulbrich, Ziems...
- P. Benner, E.W. Sachs, S. V.: *Model order reduction for PDE constrained optimization*, 2014
- C.-C. Gräßle: *POD based inexact SQP methods for optimal control problems governed by a semilinear heat equation*, 2014
- M. Kahlbacher, S.V. *POD a-posteriori error based inexact SQP method for bilinear elliptic optimal control problems*, 2012
- E. Kammann, F. Tröltzsch, S.V.: *A method of a-posteriori error estimation with application to POD*, 2013
- O. Lass, S. V.: *Parameter identification for nonlinear elliptic-parabolic systems with application in lithium-ion battery modeling*, 2014
- S. Rogg: *Trust region POD for optimal boundary control of a semilinear heat equation*, 2014
- E.W. Sachs, S.V.: *POD-Galerkin approximations in PDE-constrained optimization*, 2010
- S.V.: *Proper Orthogonal Decomposition zur Optimalsteuerung linearer partieller Differentialgleichungen*, 2013