Model order reduction for Optimal Control Problems

Kolloquium

Martin Gubisch

University of Konstanz

April 16, 2015

Universität Konstanz

Outline

1. The optimal control problem
2. Model reduction
3. Numerical experiments
4. References

Universität Konstanz

Problem formulation

We consider the optimal control problem (OCP)

$$\min_{y,u,w} J(y,u,w) = \iint\limits_{\Theta\,\Omega} \frac{1}{2}|y(t,x) - y_d(t,x)|^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{\sigma_u}{2}\|u\|^2_{L^2(\Theta,\mathbb{R}^\mathsf{m})}$$

subject to the linear parabolic pde constraint

$$\begin{aligned} \dot{y}(t,x) - \Delta y(t,x) &= (\mathcal{B}u)(t,x) && \text{in } \Theta\times\Omega,\\ y(t,x) &= 0 && \text{in } \Theta\times\partial\Omega,\\ y(0,x) &= 0 && \text{in } \Omega \end{aligned}$$

and the control and state constraints

$$y_a \le y(t,x) \le y_b \qquad \& \qquad u_a \le u(t) \le u_b,$$

Problem formulation

We consider the optimal control problem (OCP)

$$\min_{y,u,w} J(y,u,w) = \iint\limits_{\Theta\,\Omega} \frac{1}{2}|y(t,x) - y_d(t,x)|^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{\sigma_u}{2}\|u\|^2_{L^2(\Theta,\mathbb{R}^\mathsf{m})} + \frac{\sigma_w}{2}\|w(t)\|^2_{L^2(\Theta,\mathbb{R}^\mathsf{n})}$$

subject to the linear parabolic pde constraint

$$\begin{aligned} \dot{y}(t,x) - \Delta y(t,x) &= (\mathcal{B}u)(t,x) && \text{in } \Theta\times\Omega,\\ y(t,x) &= 0 && \text{in } \Theta\times\partial\Omega,\\ y(0,x) &= 0 && \text{in } \Omega \end{aligned}$$

and the control and state constraints

$$y_a \le \varepsilon w(t) + (\mathcal{I}y)(t) \le y_b \qquad \& \qquad u_a \le u(t) \le u_b,$$

with the operators $\mathcal{B} : L^2(\Theta,\mathbb{R}^\mathsf{m}) \to L^2(\Theta,H)$ and $\mathcal{I} : L^2(\Theta,H) \to L^2(\Theta,\mathbb{R}^\mathsf{n})$,

$$(\mathcal{B}u)(t,x) = \sum_{i=1}^{\mathsf{m}} u_i(t)\chi_i(x), \qquad (\mathcal{I}y)_i(t) = \int\limits_{\Omega_i} y(t,x)\,\mathrm{d}x.$$

Transformation on pure box constraints

Introducing a transformed penalty $\omega(t) = \varepsilon w(t) + \mathcal{I}y(t)$, we get the equivalent transformed optimal control problem (TOCP)

$$\min_{y,u,\omega} \tilde{J}(y,u,\omega) = \int_\Theta \frac{1}{2}\|y(t) - y_d(t)\|_H^2 \,\mathrm{d}t + \frac{\sigma_u}{2}\|u(t)\|_{\mathbb{R}^m}^2 + \frac{\sigma_w}{2\varepsilon^2}\|\omega(t) - \mathcal{I}y(t)\|_{\mathbb{R}^n}^2$$

subject to the evolution equation

$$E(y,u)(t) = \dot{y}(t) - \Delta y(t) - \mathcal{B}u(t) = 0 \text{ in } V' \qquad \& \qquad y(0) = 0 \text{ in } H$$

with $V = H_0^1(\Omega)$, $H = L^2(\Omega)$ and the *explicit* penalty and control constraints

$$\hat{y}_a \le \omega(t) \le \hat{y}_b \qquad \& \qquad u_a \le u(t) \le u_b.$$

Universität Konstanz

Well-posedness and optimality conditions

THEOREM. There exists a unique solution $(\bar{y}, \bar{u}, \bar{\omega})$ to (TOCP).

Introducing the Langange function

$$\mathscr{L}(y,u,\omega,p) = \tilde{J}(y,u,\omega) + \int_\Theta \langle E(y,u)(t), p(t)\rangle_{V',V} \,\mathrm{d}t,$$

we get the variational optimality conditions

$$\begin{aligned}
\mathscr{L}_p(\bar{y},\bar{u},\bar{\omega},\bar{p}) &= 0; \\
\mathscr{L}_y(\bar{y},\bar{u},\bar{\omega},\bar{p}) &= 0; \\
\mathscr{L}_u(\bar{y},\bar{u},\bar{\omega},\bar{p})(u - \bar{u}) &\ge 0 \qquad \forall u \in [u_a, u_b]; \\
\mathscr{L}_\omega(\bar{y},\bar{u},\bar{\omega},\bar{p})(\omega - \bar{\omega}) &\ge 0 \qquad \forall \bar{\omega} \in [\hat{y}_a, \hat{y}_b].
\end{aligned}$$

Universität Konstanz

Well-posedness and optimality conditions

With linear operators $\mathcal{L}_1, \mathcal{L}_2$ and nonlinear operators $\mathcal{N}_1, \mathcal{N}_2$, the following first-order optimality conditions are satisfied:

$$\begin{aligned} \dot{y} - \Delta y - \mathcal{B}u &= 0 \\ -\dot{p} - \Delta p - \mathcal{L}_1 y - \mathcal{L}_2 \omega &= 0 \\ u - \mathcal{N}_1(p)p &= 0 \\ \omega - \mathcal{N}_2(y)y &= 0 \end{aligned}$$

The system can be iteratively solved by the primal-dual active set strategy (PDASS)

$$\begin{aligned} \dot{y}_{k+1} - \Delta y_{k+1} &= \mathcal{B}\mathcal{N}_1(p_k)p_{k+1} \\ -\dot{p}_{k+1} - \Delta p_{k+1} &= (\mathcal{L}_1 + \mathcal{L}_2\mathcal{N}_2(y_k))y_{k+1}. \end{aligned}$$

This is a semismooth Newton method with global convergence and superlinear convergence rates.

Universität
Konstanz

Proper orthogonal decomposition (POD)

Variational form: For all test functions $\varphi \in V$ we postulate

$$\begin{aligned} \langle \dot{y} - \Delta y - \mathcal{B}\mathcal{N}_1 p, \varphi \rangle_{V',V} &= 0, \\ \langle -\dot{p} - \Delta p - (\mathcal{L}_1 + \mathcal{L}_2\mathcal{N}_2)y, \varphi \rangle_{V',V} &= 0. \end{aligned}$$

Discretization: Replace the test space V by a finite dimensional subspace $V^N \subseteq V$ spanned by $\varphi_1, ..., \varphi_N \in V$ and search approximate solutions y^N, p^N of the form

$$y^N(t,x) = \sum_{i=1}^N \mathrm{y}_i(t)\varphi_i \qquad \& \qquad p^N(t,x) = \sum_{i=1}^N \mathrm{p}_i(t)\varphi_i.$$

Problem: The discretized linear systems are still of the dimension $2N$.

Idea: For $\ell \ll N$, find an *optimal* orthonormal system $\psi = (\psi_1, ..., \psi_\ell) \subseteq V$ such that the projection error of y on the space $\mathrm{span}(\psi)$ is minimal:

$$\min_{\psi \text{ ONB}} \int_\Theta \left\| y(t) - \sum_{i=1}^{\ell} \langle y(t), \psi_i \rangle_V \psi_i \right\|_V^2 \mathrm{d}t. \tag{POD}$$

Universität
Konstanz

Proper orthogonal decomposition (POD)

THEOREM. (Continuous version) Let $y \in C(0, T; V)$ be an *arbitrary* state and let $(\lambda_i, \psi_i)_{i \in \mathbb{N}}$ be a normalized eigenvalue decomposition of the compact, nonnegative, selfadjoint operator

$$\mathcal{R}(y) : V \to V, \qquad \mathcal{R}(y)\varphi = \int_\Theta \langle y(t), \varphi \rangle_V y(t)\, \mathrm{d}t.$$

with $\lambda_i \geq \lambda_{i+1}$ for all $i \in \mathbb{N}$.

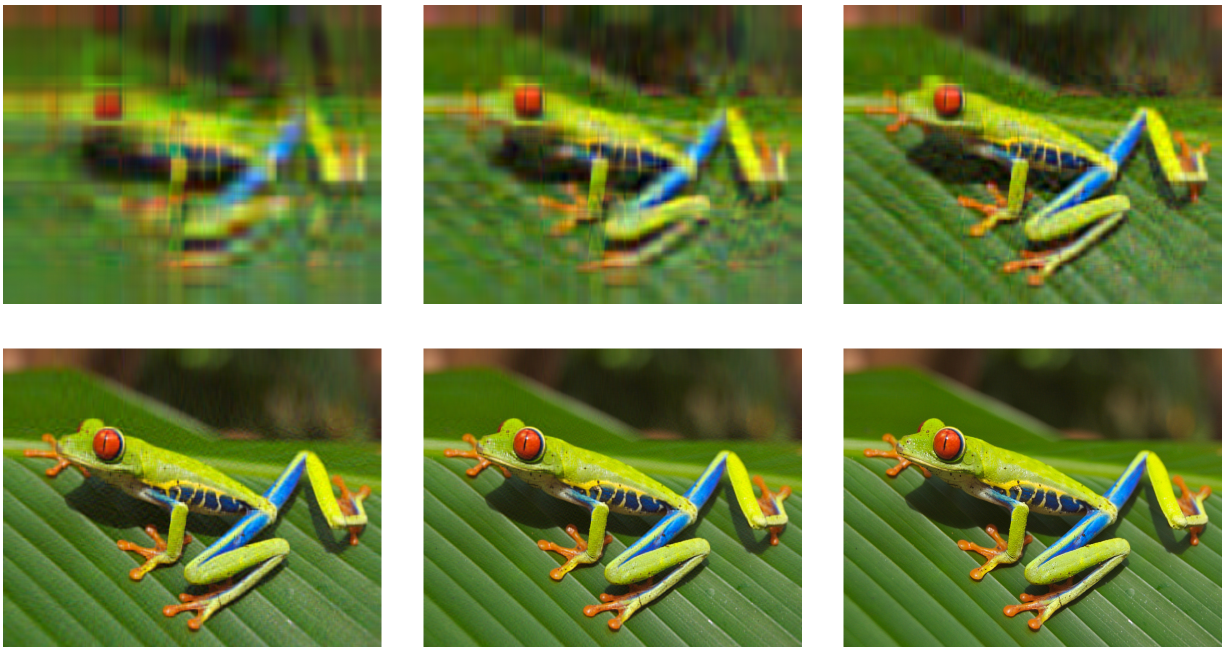
Then the *rank-ℓ POD basis* $\psi^\ell = (\psi_1, ..., \psi_\ell)$ is a solution to (POD).

A priori estimate: The projection error of y on $V^\ell = \operatorname{span}(\psi)$ fulfills

$$\int_\Theta \left\| y(t) - \sum_{i=1}^{\ell} \langle y(t), \phi_i \rangle_V \phi_i \right\|_V^2 = \sum_{i=\ell+1}^{\infty} \lambda_i.$$

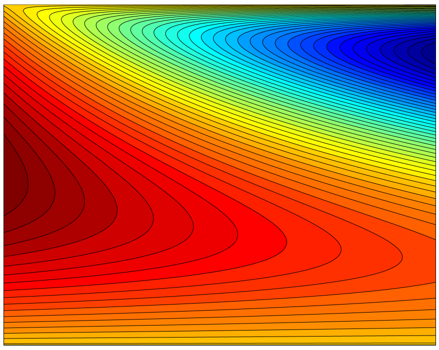
Universität Konstanz

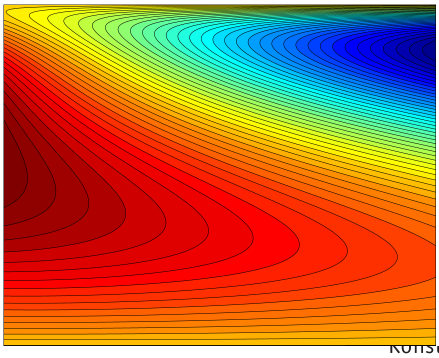
Filtering and compression of frogs



Universität Konstanz

Random data and dynamical flow



Reduced order model (ROM)

1. **ROM components:**
   - $\mathrm{M} = (\langle \psi_i, \psi_j \rangle_H) \in \mathbb{R}^{\ell \times \ell}$ and $\mathrm{A} = (\langle -\Delta \psi_i, \psi_j \rangle_{V^\star, V}) \in \mathbb{R}^{\ell \times \ell}$.
   - $\mathrm{L}_1 \in \mathrm{R}^{\ell \times \ell}$, $\mathrm{L}_2 \in \mathrm{R}^{\ell \times \mathrm{n}}$, $\mathrm{N}_1(\mathrm{p}) \in \mathrm{R}^{\mathrm{m} \times \ell}$, $\mathrm{N}_2(\mathrm{y}) \in \mathrm{R}^{\mathrm{n} \times \ell}$ and $\mathrm{B} = (\langle \mathcal{B}_j^\star, \psi_i \rangle_{V^\star, V}) \in \mathbb{R}^{\ell \times \mathrm{m}}$.

2. **ROM system:**

$$\begin{aligned} \mathrm{M}\dot{\mathrm{y}}(t) + \mathrm{A}\mathrm{y}(t) - \mathrm{B}\mathrm{N}_1(\mathrm{p})\mathrm{p}(t) &= 0, & \mathrm{y}(0) &= 0, \\ -\mathrm{M}\dot{\mathrm{p}}(t) + \mathrm{A}\mathrm{p}(t) - (\mathrm{L}_1 + \mathrm{L}_2\mathrm{N}_2(\mathrm{y}))\mathrm{y} &= 0, & \mathrm{p}(T) &= 0. \end{aligned}$$

3. **ROM expansions:**

$$y^\ell(t) = \sum_{l=1}^{\ell} \mathrm{y}_l(t)\psi_l, \quad p^\ell(t) = \sum_{l=1}^{\ell} \mathrm{p}_l(t)\psi_l,$$

$$u^\ell(t) = \mathcal{N}_1(p^\ell)p^\ell, \quad \omega^\ell(t) = \mathcal{N}_2(y^\ell)y^\ell.$$

Universität
Konstanz

Reduced order model (ROM)

A posteriori error bound: Let (u, ω) be any suboptimal control-penalty pair. Then there exists some computable $\zeta \in L^2(\Theta, \mathbb{R}^m \times \mathbb{R}^n)$ such that

$$\int_\Theta \|u(t) - \bar{u}(t)\|^2_{\mathbb{R}^m} + \|\omega(t) - \bar{\omega}(t)\|^2_{\mathbb{R}^n} \,\mathrm{d}t$$
$$\leq \frac{1}{\min(\sigma_u^2, \sigma_w^2)} \int_\Theta \|\zeta(t)\|^2_{\mathbb{R}^m \times \mathbb{R}^n} \,\mathrm{d}t + C(\Delta t + \Delta x^2).$$

Similar results are available for nonlinear PDEs; then second-order information – the smallest eigenvalue of the Hessian $\tilde{J}''$ – is required.

$\rightarrow$ SQP (Sequential Quadratic Programming) & TR-POD (Trust Region POD).

Universität
Konstanz

Reduced order model (ROM)

Algorithm (Model reduction with iterative POD basis updates (IPOD))

Require: Initial control-penalty pair $(u^{(0)}, \omega^{(0)})$, POD basis rank ℓ, desired exactness ε, maximal iteration number $k_{\max}$.

Set $k = 0$
repeat
  Solve the full state equation for $y^{(k)}$.
  Solve the POD eigenvalue problem for the rank-ℓ basis $\psi^{(k)}$.
  Provide the PDASS algorithm to solve the ROM system for $(u^{(k)}, \omega^{(k)})$.
  Set $k = k + 1$
until Aposti$(u^{(k)}, \omega^{(k)}) < \varepsilon$ or $k > k_{\max}$.
Return control $u^{(k)} \in L^2(\Theta, \mathbb{R}^m)$ and penalty $w^{(k)} = \frac{1}{\varepsilon}(\omega^{(k)} - \mathcal{I}y) \in L^2(\Theta, \mathbb{R}^n)$.

Universität
Konstanz

Optimality system proper orthogonal decomp. (OSPOD)

The optimal state required to determine the POD basis is known *implicitly*:

$$\min_{y,u,\omega,\psi} \tilde{\mathrm{J}}(\mathrm{y},u,\omega,\psi) = \int_\Theta \frac{1}{2}\Big\|\sum_{l=1}^{\ell} \mathrm{y}_l\psi_l - y_d\Big\|_H^2 + \frac{\sigma_u}{2}\|u\|_{\mathbb{R}^m}^2 + \frac{\sigma_w}{2\varepsilon^2}\|\omega - \mathrm{I}(\psi)\mathrm{y}\|_{\mathbb{R}^n}^2 \,\mathrm{d}t$$

subject to the *two* state equations

$$\dot{y} + \mathcal{A}y = \mathcal{B}u, \qquad y(0) = 0, \tag{1}$$

$$\mathrm{M}(\psi)\dot{\mathrm{y}} + \mathrm{A}(\psi)\mathrm{y} = \mathrm{B}(\psi)u, \qquad \mathrm{y}(0) = 0, \tag{2}$$

the POD eigenvalue problem

$$\mathcal{R}(y)\psi_l - \lambda_l\psi_l = 0, \qquad \|\psi_l\|_V^2 = 1 \tag{3}$$

and the penalty and control constraints

$$y_a(t) \le \omega(t) \le y_b(t) \qquad \& \qquad u_a(t) \le u(t) \le u_b(t).$$

Universität Konstanz

A priori error estimates

If POD is provided with the snapshots $y(\bar{u})$, we have the a priori error estimate

$$\|y(u) - y^\ell(u)\|_{L^2(\Theta,V)}^2 \le C(\bar{u})\Big(\sum_{l=\ell+1}^{\infty}\lambda_l + \|y_\circ - \mathcal{P}^\ell y_\circ\|_H^2 + \int_\Theta \|\dot{y}(u) - \mathcal{P}^\ell\dot{y}(u)\|_V^2\,\mathrm{d}t\Big)$$

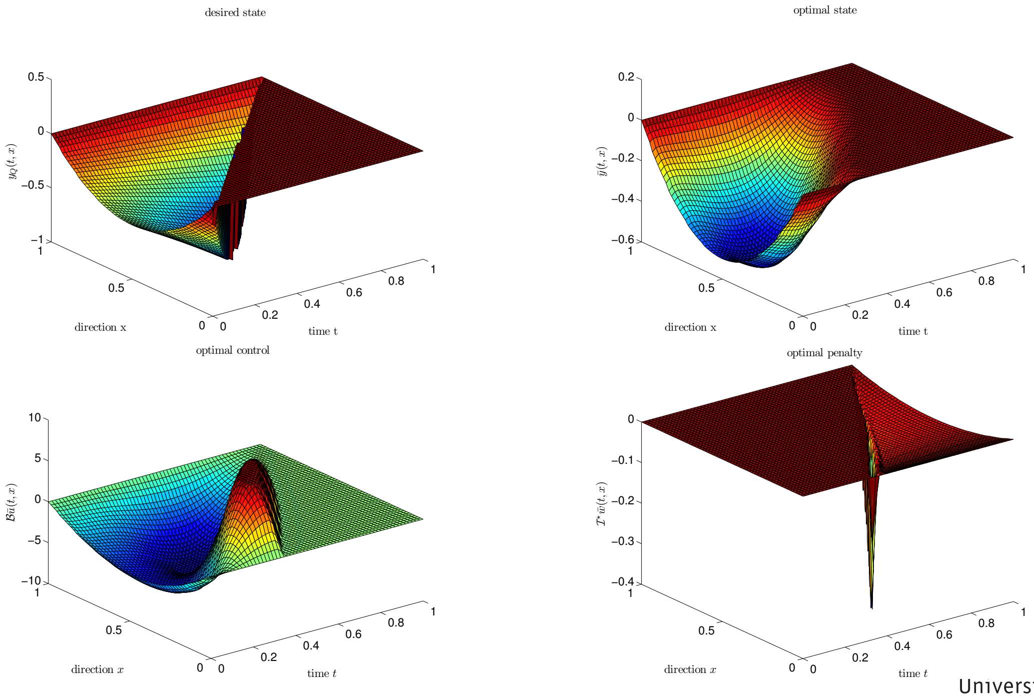
where $\mathcal{P}^\ell : V \to V^\ell$ is the projection $\phi \mapsto \sum_{l=1}^{\ell}\langle\phi,\psi_l\rangle_V\psi_l$.

With $y_\circ = 0$ and the usage of the additional snapshots $\dot{y}(\bar{u})$, OS-POD admits decay rates of the form

$$\|y(u) - y^\ell(u)\|_{L^2(\Theta,V)}^2 \le \frac{C}{\ell}\|u\|_U^2.$$

Universität Konstanz

Numerical experiments: Run 1



Universität Konstanz

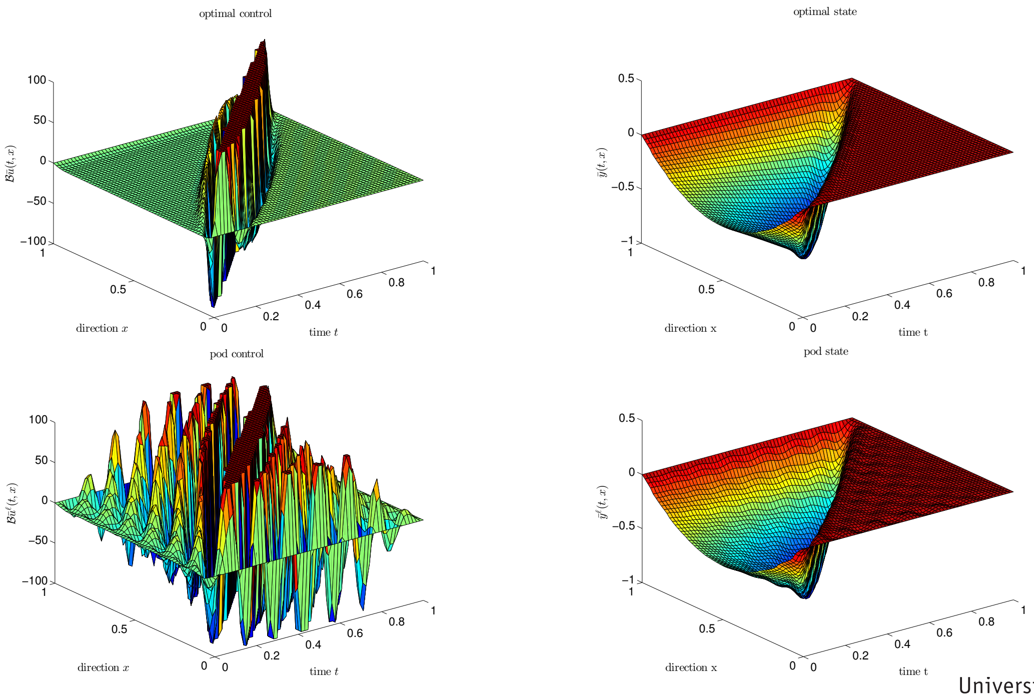
Calculation times

method	DoF	CPU time	relative time
finite element system	$N_x = 500$	860.75 sec	100.00%
initial basis	$\ell = 35$	110.77 sec	13.02%
iterative basis updates	$\ell = 15$	37.41 sec	4.40%
OS-POD basis selection	$\ell = 13$	18.39 sec	2.16%
optimal POD basis	$\ell = 13$	11.48 sec	1.35%

Universität Konstanz

Martin Gubisch (University of Konstanz) | Vortrag über wesentl. Ergebnisse (2015) | April 16, 2015 | 18 / 21

Numerical experiments: Run 2



Universität
Konstanz

References I

Arian, E., Fahl, M. & Sachs, E.: *Trust-region proper orthogonal decomposition for flow control*.
Technical Report 2000, ICASE, **vol. 25**: pp. 1–15, 2000.

Gräßle, C.: *POD based inexact SQP methods for optimal control problems governed by a semilinear heat equation*.
Master's thesis, Universität Konstanz, 2014.

Grimm, E., Gubisch, M. & Volkwein, S.: *A-Posteriori Error Analysis and OS-POD for Optimal Control*.
Comp. Eng., **vol. 3**: pp. 125–129, 2014.

Gubisch, M., Neitzel, I. & Volkwein, S.: *POD a-posteriori analysis respecting finite element discretization errors*.
to appear 2015.

Gubisch, M. & Volkwein, S.: *POD Reduced-Order Modelling for PDE Constrained Optimization*.
proceeding Model Reduction and Approximation (Luminy), submitted, Oct. 2013.
URL http://nbn-resolving.de/urn:nbn:de:bsz:352-250378.

Gubisch, M. & Volkwein, S.: *POD a-posteriori error analysis for optimal control Problems with mixed constraints*.
Comp. Opt. Appl., **vol. 58**, no. 3: pp. 619–644, 2014.

Hintermüller, M., Kopacka, I. & Volkwein, S.: *Mesh-independence and preconditioning for control problems*.
ESAIM: COVC, **vol. 15**: pp. 626–652, 2009.

Hinze, M. & Volkwein, S.: *Error estimates for abstract linear-quadratic optimal control problems using POD*.
Comput. Optim. Appl., **vol. 39**: pp. 319–345, 2008.

Kammann, E., Tröltzsch, F. & Volkwein, S.: *A-posteriori error estimation with application to POD*.
ESAIM M2AN, **vol. 47**: pp. 555–581, 2013.

Universität
Konstanz

References II

- Kunisch, K. & Müller, M.: *Uniform convergence of the Pod method and applications to optimal control.* AIMS Journals, **vol. ??**, no. ??: pp. 1–26, 2014.
- Kunisch, K. & Volkwein, S.: *Proper orthogonal decomposition for optimality systems.* ESAIM M2AN, **vol. 42**: pp. 1–23, 2008.
- Rogg, S.: *Trust Region POD for Optimal Boundary Control of a Semilinear Heat Equation.* Master's thesis, Universität Konstanz, 2014.
- Tröltzsch, F. & Volkwein, S.: *POD a-posteriori error estimates for linear-quadratic optimal control problems.* Comput. Optim. Appl., **vol. 44**: pp. 319–345, 2008.

Universität Konstanz