Suboptimal Open-loop Control Using POD

Stefan Volkwein

Institute for Mathematics and Scientific Computing
University of Graz, Austria

PhD program in Mathematics for Technology
Catania, May 22, 2007

Motivation

- Optimal control of evolution problems:

$$\min \mathrm{J}(\mathrm{y},\mathrm{u}) \quad \text{s.t.} \quad \dot{\mathrm{y}}(t) = \mathrm{F}(\mathrm{y}(t),\mathrm{u}(t)) \text{ for } t > 0,\ \mathrm{y}(0) = \mathrm{y}_\circ,\ \mathrm{u} \in \mathcal{U}$$

- Optimization methods:
  - First-order methods: gradient type methods
    $\Rightarrow$ per iteration nonlinear state and linear adjoint equations
  - Second-order methods: SQP or Newton methods
    $\Rightarrow$ per iteration coupled linear state and linear adjoint equations
- Spatial discretization by FE or FD
  $\Rightarrow$ large-scale problems and feedback-strategies not feasible
- Model reduction by POD

Outline

- Suboptimal control of a nonlinear heat equation
- Suboptimal control of the Navier-Stokes equation
- POD for parameter estimation

Nonlinear heat equation [Diwoky/V.]

- Model problem:

$$\min J(y,u) = \frac{1}{2}\int_\Omega |y(T,x)-z(x)|^2\,\mathrm{d}x + \frac{\beta}{2}\int_0^T\int_\Gamma |u(t,s)|^2\,\mathrm{d}s\mathrm{d}t$$

subject to

$$\begin{aligned} y_t(t,x) &= k\Delta y(t,x) && \text{for } (t,x)\in Q=(0,T)\times\Omega \\ \frac{\partial y}{\partial n}(t,s) &= b(y(t,s)) + u(t,s) && \text{for } (t,s)\in\Sigma=(0,T)\times\Gamma \\ y(0,x) &= y_\circ(x) && \text{for } x\in\Omega\subsetneq\mathbb{R}^2 \end{aligned}$$

- Assumptions: $T,\beta,k>0$, $z,y_\circ\in C(\overline{\Omega})$, $b\in C^{2,1}(\mathbb{R})$ with $b'\le 0$

Infinite-dimensional problem

- Optimization variables: $z = (y, u) \in Z$, Z function space
- Equality constraints: $e = (e_1, e_2)$

$$\langle e_1(z), \varphi\rangle = \int_0^T \int_\Omega y_t(t,x)\varphi(t,x) + k\nabla y(t,x)\cdot\nabla\varphi(t,x)\,\mathrm{d}x\mathrm{d}t - \int_0^T \int_\Gamma \big(b(y(t,s)) + u(t,s)\big)\varphi(t,s)\,\mathrm{d}s\mathrm{d}t$$

$$e_2(z) = y(0,\cdot) - y_\circ$$

- Infinite-dimensional optimization in function spaces:

$$\min J(z) \quad \text{subject to} \quad e(z) = 0$$

- Lagrange function: $L(z, p) = J(z) + \langle e(z), p\rangle$
- Optimality conditions: $\nabla L(z, p) \overset{!}{=} 0$ (Fréchet-derivatives)

First-order optimality conditions

- $\nabla_y L(y,u,p) \stackrel{!}{=} 0$: adjoint equation

$$
\begin{aligned}
-p_t(t,x) &= k\Delta p(t,x) && \text{for } (t,x) \in Q = (0,T)\times\Omega\\
\frac{\partial p}{\partial n}(t,s) &= b'(y(t,s))p(t,s) && \text{for } (t,s) \in \Sigma = (0,T)\times\Gamma\\
p(T,x) &= -(y(T,x) - z(x)) && \text{for } x \in \Omega
\end{aligned}
$$

- $\nabla_u L(z,p) \stackrel{!}{=} 0$: optimality condition $\beta u = kp$ on Σ
- $\nabla_p L(z,p) \stackrel{!}{=} 0$: state equation

$$
\begin{aligned}
y_t(t,x) &= k\Delta y(t,x) && \text{for } (t,x) \in Q\\
\frac{\partial y}{\partial n}(t,s) &= b(y(t,s)) + u(t,s) && \text{for } (t,s) \in \Sigma\\
y(0,x) &= y_\circ(x) && \text{for } x \in \Omega
\end{aligned}
$$

SQP methods

- SQP: sequentiel quadratic programming
- Quadratic programming problem: $L(z,p) = J(z) + \langle e(z), p \rangle$

$$\begin{aligned} &\min L(z^n,p^n) + L_z(z^n,p^n)\delta z + \tfrac{1}{2} L_{zz}(z^n,p^n)(\delta z, \delta z) \\ &\text{subject to } e(z^n) + e'(z^n)\delta z = 0 \end{aligned} \qquad (\mathrm{QP}^n)$$

- First-order optimality conditions for (QP^n): KKT system

$$\begin{pmatrix} L_{zz}(z^n,p^n) & e'(z^n)^\star \\ e'(z^n) & 0 \end{pmatrix} \begin{pmatrix} \delta z \\ \delta p \end{pmatrix} = - \begin{pmatrix} L_z(z^n,p^n) \\ e(z^n) \end{pmatrix}$$

- Convergence: locally quadratic rate in (z^n,p^n) (infinite-dimensional)
- Globalization: modification of the Hessian and line-search methods
- Alternative: trust-region methods

POD model reduction

- Goal: POD Galerkin ansatz using ℓ POD basis functions
- Snapshot POD: solve of heat equation for $0 \leq t_1 < \ldots < t_n \leq T$
- Problems:
  - unknown optimal control $\Rightarrow$ good snapshot set?
  - $u = \frac{k}{\beta} p$ depends on p $\Rightarrow$ POD approximation for p?
- Strategy: iterate basis computation and include adjoint information in the snapshot ensemble

Dynamic POD strategy [Hinze et al./Sachs et al.]

(1) Choose estimate u^0; compute snapshots by solving state equation with $u = u^0$ and adjoint equation with $y = y(u^0)$; $i := 0$

(2) Determine ℓ POD basis functions and associated ROM of infinite-dimensional optimization problem

(3) Compute solution u^{i+1} of optimization problem (e.g., by SQP)

(4) If $\Psi(i) = \dfrac{\|u^{i+1} - u^i\|}{\|u^{i+1}\|} \leq TOL$ then stop (stopping criterium)

(5) $i := i + 1$; compute snapshots by solving state equation with control $u = u^i$ and adjoint equation with $y = y(u^i)$; go back to (2)

Numerical results

Data: $y_\circ(x_1, x_2) = 10x_1x_2$, $z(x_1, x_2) = 2 + 2|2x_1 - x_2|$, $b(y) = \arctan(y)$, $k = \beta = \frac{1}{10}$, $T = 1$, 185 FEs

Recall: $\Psi(i) = \frac{\| u^{i+1} - u^i \|}{\| u^{i+1} \|}$ stopping criterium for dynamic POD strategy

i	relative L^2 error for y	relative L^2 error for u	$J(y,u)$	$\Psi(i)$
0	4.4	12.0	0.358	1.00
1	1.0	8.1	0.360	0.13
2	0.9	6.8	0.361	0.08
POD_{opt}	0.5	5.7	0.358	
FE			0.358	

		POD	FE
Compute snapshots	M-flops	18	
	CPU time in s	3.3	
Compute POD basis	M-flops	0.44	
	CPU time in s	0.01	
Solve with SQP	M-flops	84	
	CPU time in s	22	
total	M-flops	$1.0 \cdot 10^2$	$1.9 \cdot 10^5$
	CPU time in s	$2.5 \cdot 10^1$	$6.6 \cdot 10^3$

Suboptimal control

PDE:

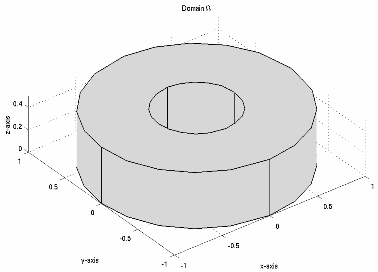
$$
\begin{aligned}
y_t &= \Delta y && \text{in } (0,T)\times\Omega \\
\frac{\partial y}{\partial n} &= 0 && \text{on } (0,1)\times\Gamma_1 \\
\frac{\partial y}{\partial n} &= u(t)q && \text{on } (0,1)\times\Gamma_2 \\
y(0) &= 0 && \text{on } \Omega\subset\mathbb{R}^3
\end{aligned}
$$

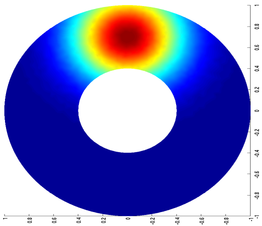
Boundary condition at $z = 0.5$:

$$
q(t,\mathrm{x}) = e^{-\left(x-0.7\cos(2\pi t)\right)^2} \cdot e^{-\left(y-0.7\sin(2\pi t)\right)^2}
$$

Cost functional:

$$
J(y,u) = \frac{1}{2}\int_\Omega |y(T)-1|^2\,\mathrm{d}x + \frac{\sigma}{2}\int_0^T |u(t)|^2\,\mathrm{d}t
$$





POD computation

- Snapshot ensembles:

$$\mathcal{V}_1 = \operatorname{span}\left\{\{\bar{y}^h(t_j)\}_j, \left\{\frac{\bar{y}^h(t_j)-\bar{y}^h(t_{j-1})}{\Delta t}\right\}_j\right\}$$

$$\mathcal{V}_2 = \operatorname{span}\left\{\{\bar{p}^h(t_j)\}_j, \left\{\frac{\bar{p}^h(t_j)-\bar{p}^h(t_{j-1})}{\Delta t}\right\}_j\right\}$$

$$\mathcal{V}_3 = \mathcal{V}_1 \cup \mathcal{V}_2$$

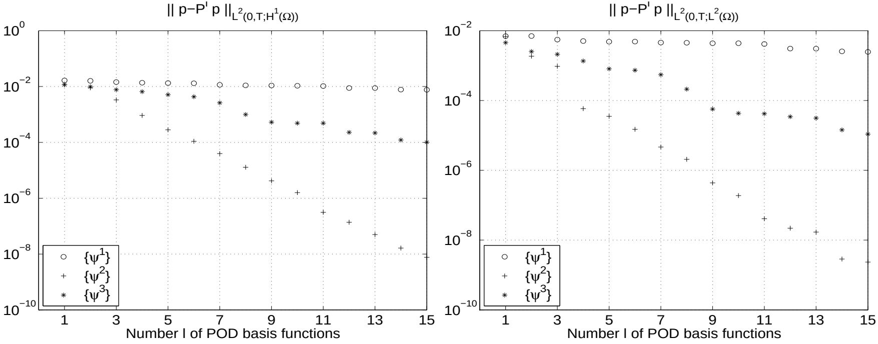
- $E(\ell) = \sum_{i=1}^{\ell} \lambda_i \cdot 100\%$:

ℓ	$E(\ell)$ for $\mathcal{V}^1$	$E(\ell)$ for $\mathcal{V}^2$	$E(\ell)$ for $\mathcal{V}^3$
$\ell = 1$	45.89 %	70.44 %	48.20 %
$\ell = 3$	87.65 %	97.41 %	84.39 %
$\ell = 7$	99.37 %	100.00 %	98.06 %
$\ell = 11$	99.78 %	100.00 %	99.82 %
$\ell = 15$	99.80 %	100.00 %	99.90 %

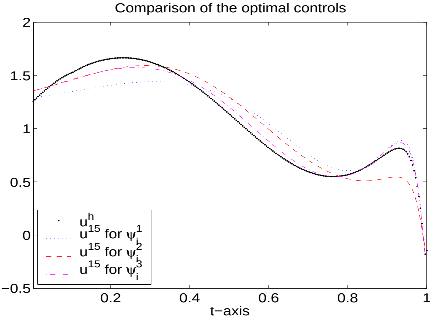
Approximation of the dual variable

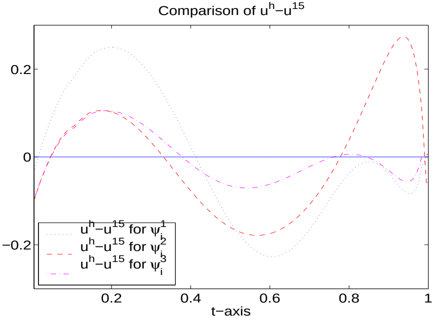
Theoretical POD estimate for the dual:

$$\|p^\ell - p\|_{L^2(0,T;V)} \leq C\Big(\|y^\ell - y\| + \|p - \mathcal{P}^\ell p\|\Big)$$



Approximation of the control variable





ℓ	$\|u^h - u^\ell\|$ for $\{\psi_i^1\}_{i=1}^\ell$	$\|u^h - u^\ell\|$ for $\{\psi_i^2\}_{i=1}^\ell$	$\|u^h - u^\ell\|$ for $\{\psi_i^3\}_{i=1}^\ell$
$\ell = 1$	0.5100	0.5437	0.4672
$\ell = 3$	0.3792	0.1200	0.1869
$\ell = 5$	0.3506	0.0588	0.1201
$\ell = 9$	0.3031	0.0585	0.0566
$\ell = 13$	0.2057	0.0596	0.0555

$\|u^h - u^\ell\|$ for different POD basis $\{\psi_i^j\}_{i=1}^\ell$ corresponding to the ensembles $\mathcal{V}_j$, $j = 1, 2, 3$

Control of the Navier-Stokes equation [Hinze]

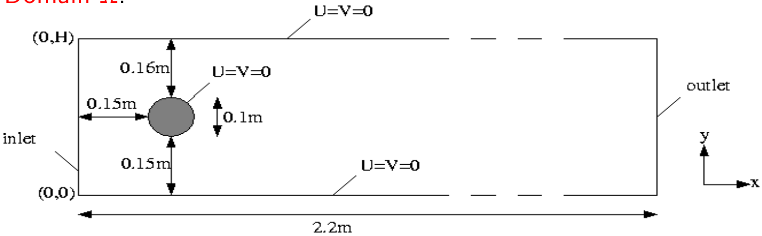
- **Optimal control problem**:

$$\min_{(y,u)} J(y,u) = \frac{1}{2}\int_0^T \int_\Omega |y-z|^2 \,\mathrm{d}x\mathrm{d}t + \frac{\alpha}{2}\int_0^T \int_\Omega |u|^2 \,\mathrm{d}x\mathrm{d}t$$

subject to

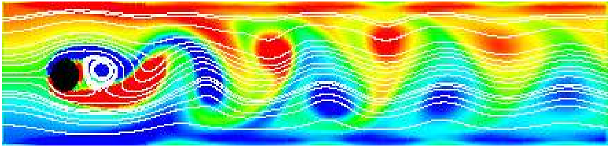
$$\begin{aligned} & y_t + (y\cdot\nabla)y - \nu\Delta y + \nabla p = \mathcal{B}u && \text{in } Q = (0,T)\times\Omega \\ & \operatorname{div} y = 0 && \text{in } Q \\ & y(t,\cdot) = y_d && \text{on } (0,T)\times\Gamma_d \\ & \nu\partial_\eta y(t,\cdot) = p\eta && \text{on } (0,T)\times\Gamma_{outlet} \\ & y(0,\cdot) = y_\circ && \text{in } \Omega \end{aligned}$$

- **Domain Ω**:

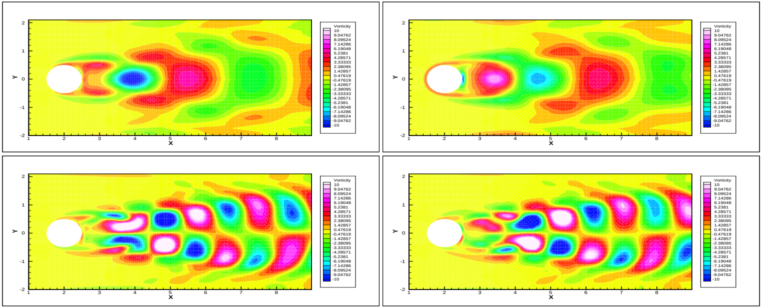


Computation of the POD basis

- Uncontrolled flow:

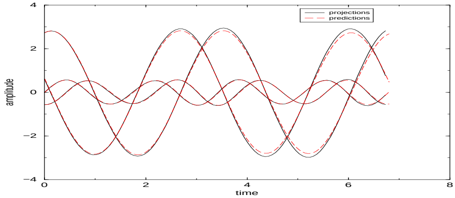


- POD basis functions: ψ_1, ψ_2, ψ_3 and ψ_4

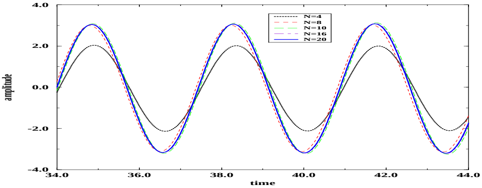


POD properties

- POD ansatz: $y^\ell = y_\mathrm{m} + \sum_{i=1}^{\ell} \alpha_i(t)\psi_i$ with $y_\mathrm{m} = \frac{1}{m}\sum_{i=1}^{m} y(t_i)$
- Projection onto the POD subspace: $\alpha_i(t)$ and $\big(y(t) - y_\mathrm{m}\big)^T \psi_i$

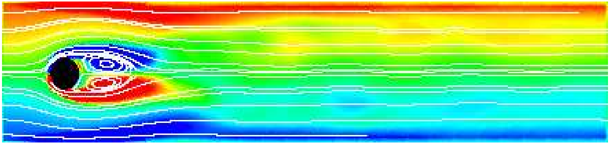


- Different number of snapshots:

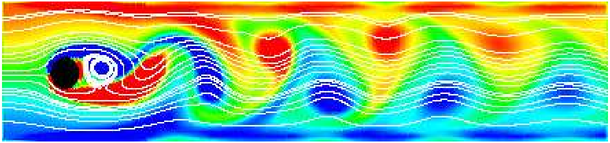


Optimal solution

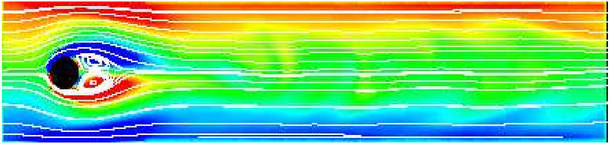
- Goal: mean flow field y_{m}



- Starting value for optimizer: uncontrolled flow, i.e., y for $u = 0$



- Optimal solution: flow field y^{ℓ}



Parameter estimation [Kahlbacher/V.]

- Model equations: $g(x) = x_1$

$$-\frac{3}{4}\,\Delta u + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \nabla u + au = g \quad \text{in } \Omega = (0,1)\times(0,1)$$
$$\frac{3}{4}\,\frac{\partial u}{\partial n} + \frac{3}{2}\,u = -1 \quad \text{on } \Gamma \qquad (*)$$

- Data: choose $a_{\mathrm{id}} \geq 0$ and compute (FE) solution $u(a_{\mathrm{id}})$ to $(*)$
- Reconstruction: estimate $a \geq 0$ from $u_d = (1+\varepsilon\delta)u(a_{\mathrm{id}})\big|_\Gamma$ with random $|\varepsilon| \leq 1$ and factor $\delta = 5\%$
- Constrained optimization:

$$\min J(a,u) = \int_\Gamma \alpha\,|u - u_d|^2\,\mathrm{d}s + \kappa\,|a|^2 \text{ s.t. } (a,u) \text{ solves } (*) \text{ and } a \geq 0$$

- Relaxation of the inequality:

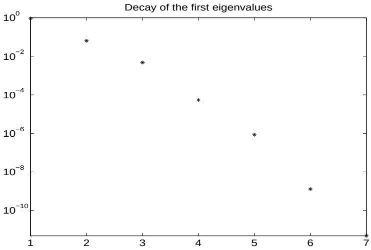
$$\min J_\lambda^\varrho(a,u) = J(a,u) + \frac{1}{\varrho}\max\left\{0, \lambda + \varrho(0-a)\right\}^2 \text{ s.t. } (a,u) \text{ solves } (*)$$

Global convergent optimization method

- Outer loop: augmented Lagrangian method → control of ϱ^k and λ^k
- Inner loop: globalized SQP algorithm with fixed (ϱ^k, λ^k) for

$$\min J^{\varrho^k}_{\lambda^k}(a,u) \quad \text{s.t.} \quad \begin{cases} -\frac{3}{4}\Delta u + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \nabla u + au = g & \text{in } \Omega \\ \frac{3}{4}\frac{\partial u}{\partial n} + \frac{3}{2}u = -1 & \text{on } \Gamma \end{cases}$$

- Numerical results: $\alpha = 5000$, $\kappa = 0.0005$, $a_{\rm id} = 25$, $\ell = 7$



relative errors:

$$\frac{\|u^\ell - u(a_{\rm id})\|}{\|u(a_{\rm id})\|} \approx 1.87 \cdot 10^{-5}$$

$$\frac{|a^\ell - a_{\rm id}|}{|a_{\rm id}|} \approx 6 \cdot 10^{-3}$$

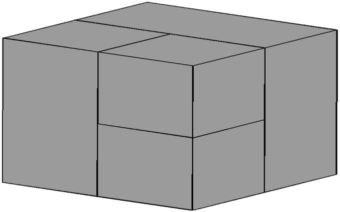
Parameter identification

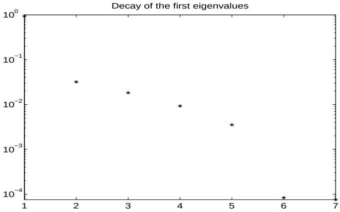
- Model equations: $a(x)$ is piecewise constant on 4 subdomains of Ω

$$-\frac{3}{4}\,\Delta u + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \nabla u + a u = 1 \quad \text{in } \Omega = (0,1)\times(0,1)\times(0,1)$$

$$\frac{3}{4}\,\frac{\partial u}{\partial n} + \frac{3}{2}\,u = 1 \quad \text{on } \Gamma \qquad (*)$$

- Data: choose $a_{\mathrm{id}} \geq 0$ and compute (FE) solution $u(a_{\mathrm{id}})$ to $(*)$
- Reconstruction: estimate $a \geq 0$ from $u_d = (1+\varepsilon\delta)u(a_{\mathrm{id}})\big|_{\Gamma}$ with random $|\varepsilon| \leq 1$ and factor $\delta = 5\%$





Global convergent optimization method

- Outer loop: augmented Lagrangian method $\rightarrow$ control of ϱ^k and λ^k
- Inner loop: globalized SQP algorithm with fixed (ϱ^k, λ^k) for

$$\min J_{\lambda^k}^{\varrho^k}(a,u) \quad \text{s.t.} \quad \begin{cases} -\frac{3}{4}\Delta u + \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \cdot \nabla u + au & = 1 \quad \text{in } \Omega \\ \frac{3}{4}\frac{\partial u}{\partial n} + \frac{3}{2}u & = 1 \quad \text{on } \Gamma \end{cases}$$

- Numerical results: $\alpha = 10000$, $\kappa = 10^{-3} \cdot (\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8})^T$, $a_{\mathrm{id}} = (10, 20, 15, 17)^T$, $\ell = 7$

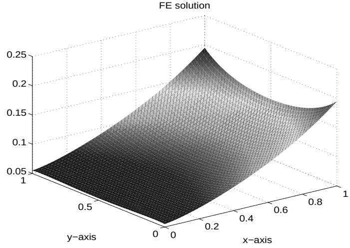
	POD	FE
$\frac{\lVert a^* - a_{\mathrm{id}} \rVert}{\lVert a_{\mathrm{id}} \rVert}$	$1.0 \cdot 10^{-2}$	$6.1 \cdot 10^{-3}$
$\frac{\lVert u^* - u(a_{\mathrm{id}}) \rVert}{\lVert u(a_{\mathrm{id}}) \rVert}$	$5.6 \cdot 10^{-3}$	$1.9 \cdot 10^{-3}$
CPU time in s	30.4	504.2

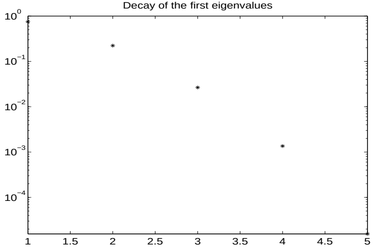
Parameter identification of the diffusion coefficient

- Model equations: $g(x) = x_1$

$$\begin{aligned} -c\,\Delta u + 25u &= 1 \quad \text{in } \Omega = (0,1)\times(0,1) \\ c\,\frac{\partial u}{\partial n} + \frac{3}{2}\,u &= g \quad \text{on } \Gamma \end{aligned} \qquad (*)$$

- Data: choose $c_{\mathrm{id}} \geq 0$ and compute (FE) solution $u(a_{\mathrm{id}})$ to $(*)$
- Reconstruction: estimate $c \geq 0$ from $u_d = (1+\varepsilon\delta)u(c_{\mathrm{id}})\big|_\Gamma$ with random $|\varepsilon| \leq 1$ and factor $\delta = 5\%$
- Optimal state and decay of the first eigenvalues:





Global convergent optimization method

- Outer loop: augmented Lagrangian method $\rightarrow$ control of ϱ^k and λ^k
- Inner loop: globalized SQP algorithm with fixed (ϱ^k, λ^k) for

$$\min J_{\lambda^k}^{\varrho^k}(c, u) \quad \text{s.t.} \quad \begin{cases} -c\,\Delta u + 25u &= 1 \quad \text{in } \Omega \\ c\,\frac{\partial u}{\partial n} + \frac{3}{2}\,u &= g \quad \text{on } \Gamma \end{cases}$$

- Numerical results: $\alpha = 100000$, $\kappa = 0.00001$, $c_{\mathrm{id}} = 3$, $\ell = 5$

	POD	FE
$\frac{\lvert c^* - c_{\mathrm{id}}\rvert}{\lvert c_{\mathrm{id}}\rvert}$	$4.17 \cdot 10^{-3}$	$4.2 \cdot 10^{-3}$
$\frac{\lVert u^* - u(c_{\mathrm{id}})\rVert}{\lVert u(c_{\mathrm{id}})\rVert}$	$6.74 \cdot 10^{-4}$	$6.72 \cdot 10^{-4}$
CPU time in s	30.4	262.2