Mathematical aspects of proper orthogonal decomposition

Lecture I: POD for time-dependent PDEs (emphasis on numerical analysis)
Lecture II: POD in PDE constrained optimization (with error analysis)
Lecture III: MOR in applications - towards parametric MOR for nonlinear PDE systems in networks
Pilsen, September 7 & 8, 2015

Michael Hinze

Lothar Collatz Center for Computing in Science

DFG SPP 1253 Optimization with partial differential equations

Collaboration

Konstantin Afanasiev, Oke Alff, Allesandro Alla, Carmen Gräßle, Christian Kahle, Martin Kunkel, Ullrich Matthes, Morten Vierling & Stefan Volkwein.

Outline

- Lecture 1: Mathematical aspects of POD
  - Motivation
  - Why Model Order Reduction?
  - Proper Orthogonal Decomposition (POD)
  - POD for Time and/or parameter dependent PDEs
  - Error estimates
  - Treatment of nonlinearities $\rightarrow$ DEIM
  - Further issues with POD
- Lecture 2: Optimization with POD surrogate models
  - Basic approach in PDE constrained optimization
  - Input dependence of POD model $\rightarrow$ POD basis updates
  - Snapshot choice in optimal control
  - Numerical analysis of POD in PDE constrained optimization
  - Further aspects of POD in applications
- Lecture 3: Towards parametric MOR for nonlinear PDE systems in networks

Motivation

We have a validated mathematical model for physical process (here a pde system)

We intend to use this model to tailor and/or optimize the physical process.

This might be computationally very expensive!

Motivation: ∞-dimensional optimization problem with pde constraints

$$\min_{(y,u)\in W\times U_{ad}} J(y,u)$$

s.t.

$$\begin{aligned} \frac{\partial y}{\partial t} + \mathcal{A}y + \mathcal{G}(y) &= \mathcal{B}u \text{ in } Z^* \\ y(0) &= y_0 \text{ in } H. \end{aligned}$$

Central tasks:

- Develop solution strategies which obey the rule

  Effort of optimization = K $\times$ Effort of simulation

  with K small,
- Propose surrogate models for the pde and quantify their errors,
- Present a complete (numerical) analysis.

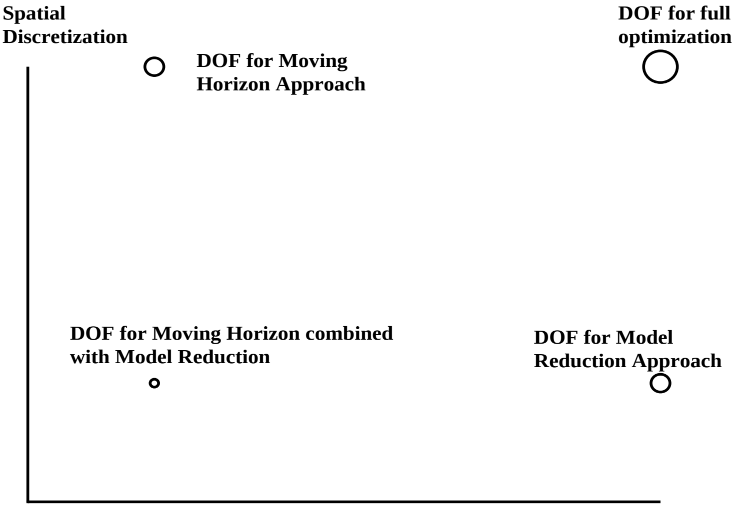
Examples of pde systems

Find $y \in W(0,T) = \{v \in L^2(0,T;V), y_t \in L^2(0,T;V^*)\}$ which solves

$$\begin{aligned} \frac{\partial y}{\partial t} + \mathcal{A}y + \mathcal{G}(y) &= \mathcal{B}u \text{ in } Z(= L^2(0,T;V)) \\ y(0) &= y_0 \text{ in } H. \end{aligned}$$

1. Heat equation: $\mathcal{A} := -\Delta$.
2. Burgers: $\mathcal{A} := -\Delta$, $\mathcal{G}(y) := yy'$,
3. Ignition (Bratu): $\mathcal{A} := -\Delta$, $\mathcal{G}(y) := -\delta e^y$, $\delta > 0$,
4. Navier–Stokes: $\mathcal{A} := -P\Delta$, $\mathcal{G}(y) := P[(y\nabla)y]$, P Leray projector,
5. Boussinesq Approximation:
   $\mathcal{A} := \begin{bmatrix} -P\Delta & -Gr\vec{g} \\ 0 & -\Delta \end{bmatrix}$, $\mathcal{G}(y) = \mathcal{G}(v,\theta) := \begin{bmatrix} P[(v\nabla)v] \\ (v\nabla)\theta \end{bmatrix}$.

DOF diagram


Spatial Discretization
DOF for full optimization
DOF for Moving Horizon Approach
DOF for Moving Horizon combined with Model Reduction
DOF for Model Reduction Approach

Motivation: parametrized PDEs

Consider for $\mu = (\mu_1, \mu_2) > 0$

$$-\text{div } (A(x;\mu)\nabla y) = f \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega,$$

with

$$A(x;\mu) = \begin{cases} \mu_1, & x \in R, \\ \mu_2, & x \in \Omega \setminus R. \end{cases}$$

Aim: find a surrogate model

$$-\text{div } (\tilde{A}\nabla y) = f \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega,$$

which represents the parameter dependent problem sufficiently well over the parameter domain.

$\rightarrow$ question will be touched in lecture III.

The beginning of snapshot POD with Sirovich '87: MOR in flow control

Navier-Stokes equations

$$\begin{aligned} \frac{\partial y}{\partial t} + (y \cdot \nabla)y - \nu \Delta y + \nabla p &= f && \text{in } Q = (0, T) \times \Omega, \\ -\text{div } y &= 0 && \text{in } Q, \\ y(t, \cdot) &= g && \text{on } \Sigma = (0, T) \times \partial\Omega, \\ y(0, \cdot) &= y_0 && \text{in } \Omega. \end{aligned}$$

Aim: Reduced description of the Navier-Stokes equations

$$\begin{aligned} \dot{\alpha} + A\alpha + n(\alpha) &= r && \text{in } (0, T) \\ \alpha(0) &= a_0 \end{aligned}$$

1. Construction and validation of the reduced model

System reduction: Expansions w.r.t. base flows

Let $\bar{y}$ denote a base flow and Φ^i, $i = 1, \ldots, n$ Modes.
Ansatz for the flow:

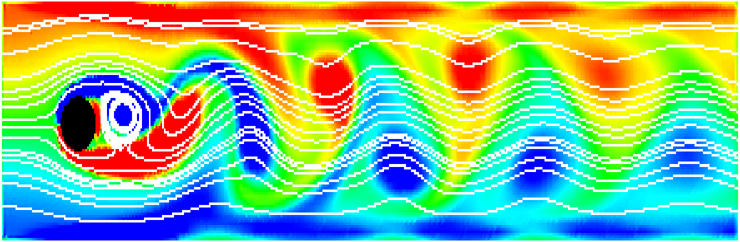
$$y = \bar{y} + \sum_{i=1}^{n} \alpha_i \Phi^i$$

Possibilities:

- **$\bar{y}$ stationary solution of Navier-Stokes system, Φ^i eigenfunctions of the Navier-Stokes system linearized at $\bar{y}$.**
- **$\bar{y}$ mean value of instationary Navier-Stokes solution, Φ^i eigenfunctions of the Navier-Stokes system linearized at $\bar{y}$.**
- **$\bar{y}$ mean value of instationary Navier-Stokes solution, Φ^i normalized Modes obtained from snapshot form of Proper Orthogonal Decomposition.**

Snapshot form of POD

Let's take snapshots:



POD with Snapshots

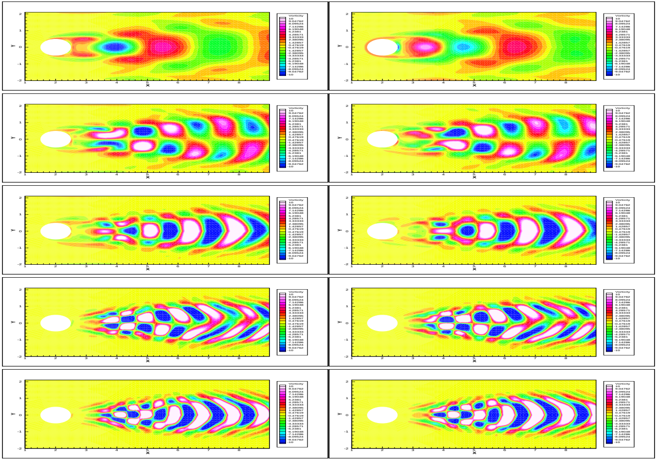
Let $y^1, \ldots, y^n$ denote an ensemble of snapshots (of the flow or the dynamical system). Build mean $\bar{y}$ and modes Φ_i as follows:

1. Compute mean $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y^i$
2. Build correlation matrix $K = k_{ij}$, $k_{ij} = \langle y^i - \bar{y}, y^j - \bar{y} \rangle$
3. Compute eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $v^1, \ldots, v^n$ of K
4. Define modes $\Phi^i := \sum_{j=1}^{n} v^i_j (y^j - \bar{y})$
5. Normalize modes $\Phi^i = \frac{\Phi^i}{\|\Phi^i\|}$

Properties:

- The modes are pairwise orthogonal w.r.t. inner product $\langle \bullet, \bullet \rangle$
- No other basis can contain more information in fewer elements (Information w.r.t. the norm induced by $\langle \bullet, \bullet \rangle$).

First 10 Modes containing 99.99 % of the information



Galerkin projection

Ansatz for the flow

$$y = \bar{y} + \sum_{i=1}^{n} \alpha_i \Phi^i$$

Galerkin method with basis $\Phi_1, \ldots, \Phi_n$ yields reduced system

$$\dot{\alpha} + A\alpha + n(\alpha) = r \quad \alpha(0) = a_0.$$

Here, $\langle \bullet, \bullet \rangle$ denotes the L^2 inner product.

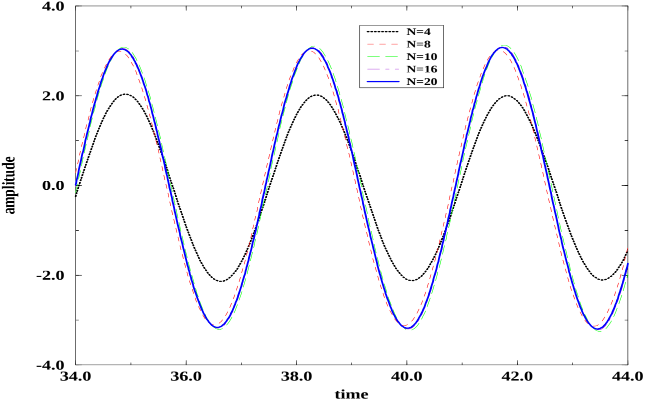
$$A = (a_{i,j})_{i=1}^{n}, \quad a_{i,j} = \nu \int_{\Omega} \nabla\Phi_i \nabla\Phi_j \, dx, \quad n(\alpha) = \left(\int_{\Omega_c} (y\nabla y)\Phi_i \, dx \right)_{i=1,}^{n}$$

$$r = -\nu \left(\int_{\Omega} \nabla\bar{y}\nabla\Phi_i + f\Phi_i \, dx \right)_{i=1}^{n} \text{ and } a_0 = \left(\int_{\Omega} y_0 \Phi_i \, dx \right)_{i=1.}^{n}$$

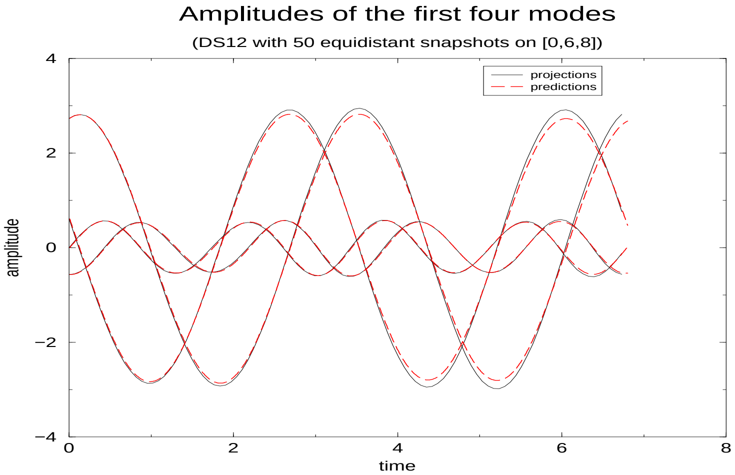
Note that $\Phi_1, \ldots, \Phi_n$ are solenoidal.

Long–time behaviour of the POD model

Amplitudes of the first mode in [34,44] when using N modes in the POD model



Cylinder flow at Re = 100, reduced versus full model, 50 snapshots



What did Sirovich propose?

- **Take snapshots $y(t_1), \dots, y(t_n)$,**
- **perform a singular value decomposition with**

$$Y := [y(t_1), \dots, y(t_n)] = \Phi \Sigma V^t,$$

where $\Sigma = diag(\sqrt{\lambda_i})$,

- **perform a Galerkin method with those modes $\Phi_1, \dots, \Phi_l$ as basis elements which carry as much information as required (say 99%, say).**

Todays point of view

Find a basis $\Phi_1, \ldots, \Phi_l \in V$ such that

$$\{\Phi_1, \ldots, \Phi_l\} = \arg\min \int_0^T \|y(t) - \sum_{i=1}^{l} \langle y(t), \Phi_i \rangle \Phi_i\|_V^2 \, dt.$$

On the discrete level we solve instead ($y(t_j)$ are $N-$vectors, so are Φ_j)

$$\min_{\Phi_1,\ldots,\Phi_l} \sum_{j=0}^{n} \beta_j \left\| y(t_j) - \sum_{i=1}^{l} \langle y(t_j), \Phi_i \rangle \Phi_i \right\|^2$$
$$\text{s.t. } \langle \Phi_j, \Phi_i \rangle = \delta_{ij} \quad \text{for } 1 \leq i,j \leq l,$$

where β_j are nonnegative quadrature weights for $\int_0^T \cdot dt$.

The projection error then satisfies

$$\sum_{j=0}^{n} \beta_j \left\| y(t_j) - \sum_{i=1}^{l} \langle y(t_j), \Phi_i \rangle \Phi_i \right\|^2 = \sum_{i=l+1}^{n} \lambda_i.$$

Some remarks

- **The choice of the snapshots is very important.**
- **Generation of snapshots with time-adaptivity.**
- **Snapshots should comply with physical properties of the underlying dynamical system, like periodicity of the flow, say.**
- **The Galerkin basis depends on the input (initial state y_0, rhs $\mathcal{B}u$).**

Error estimate (Kunisch and Volkwein (Numer. Math. 2001, SINUM 2002))

The error analysis for POD reduced systems is now along the lines of error analysis for Galerkin approximations of time dependent problems;

Let $y(t_1), \ldots, y(t_n)$ denote snapshots taken on an equidistant time grid of $[0, T]$ with gridsize δt. Let $\lambda_1 > \cdots > \lambda_d > 0$ denote the strictly positive eigenvalues of the correlation matrix K. For $l \le d$ let $V_l = \langle \Phi_1, \ldots, \Phi_l \rangle$. Further set

$$Y_k := \sum_{i=1}^{l} \alpha_i(t_k)\Phi_i .$$

Then

$$\delta t \sum_{i=1}^{n} |Y_i - y(t_i)|_H^2 \le C \left\{ \sum_{i=l+1}^{d} |\langle y_0, \Phi_i \rangle_V|^2 + \frac{1}{\delta t^2} \sum_{i=l+1}^{d} \lambda_i + \delta t^2 \right\} .$$

- This result also extends to the case of distinguish time and snapshot grids.
- Improvements of reduced models and error estimate by different weighting of snapshots (include derivative information).

Wave equations (Herkt, H. Pinnau, ETNA 2013)

Let $V \hookrightarrow H = H' \hookrightarrow V'$ denote a Gelfand triple. Consider the linear wave equation

$$\begin{aligned} \langle \ddot{x}(t), \phi\rangle_H + D\langle \dot{x}(t), \phi\rangle_H + a(x(t), \phi) &= \langle f(t), \phi\rangle_H \\ & \quad \text{for all } \phi \in V \text{ and } t \in [0, T], \\ \langle x(0), \psi\rangle &= \langle x_0, \psi\rangle_H \quad \text{for all } \psi \in H, \\ \langle \dot{x}(0), \psi\rangle &= \langle \dot{x}_0, \psi\rangle_H \quad \text{for all } \psi \in H, \end{aligned}$$

Then POD based on the Newark scheme delivers an error estimate of the form

$$\begin{aligned} \Delta t \sum_{k=1}^{m} \left\| X^k - x(t_k)\right\|_H^2 &\leq \\ \leq \quad & C_l \Bigg(\left\| X^0 - P^l x(t_0)\right\|_H^2 + \left\| X^1 - P^l x(t_1)\right\|_H^2 + \Delta t \left\| \partial X^0 - P^l \dot{x}(t_0)\right\|_H^2 \\ & + \Delta t \left\| \partial X^1 - P^l \dot{x}(t_1)\right\|_H^2 + \Delta t^4 + \left(\frac{1}{\Delta t^4} + \frac{1}{\Delta t} + 1\right) \sum_{j=l+1}^{d} \lambda_{lj} \Bigg) \end{aligned}$$

- **In general only linear decay of modes.**
- **Critical dependence on Δt can be avoided by including derivative information into the snapshot set.**

Decay of singular values for POD with parabolic equations

Linear heat equation with $y_0 \equiv 0$ and inhomogeneous boundary data.
FE-solution $\{y^h(t_j)\}_{j=0}^m$ computed on equi–distant time grid.
Snapshots:

$$y_j = \begin{cases} y^h(t_{j-1}) & \text{for } 1 \leq j \leq m+1, \\ \dfrac{y^h(t_{j-m-1}) - y^h(t_{j-m-2})}{\Delta t} & \text{for } m+2 \leq j \leq 2m+1. \end{cases}$$

Correlation matrix

$$(k_{ij})_{i,j=1}^{2m+1}, \;\; k_{ij} = \langle y_i, y_j \rangle_V$$

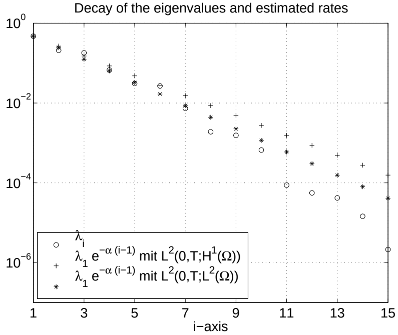
Expected decay of its eigenvalues:

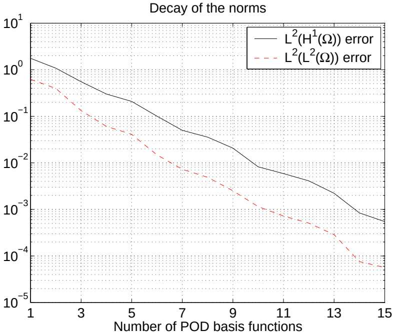
$$\lambda_i = \lambda_1 e^{-\alpha(i-1)} \quad \text{for } i \geq 1.$$

Experimental order of decay:

$$Q(\ell) = \ln \frac{\|y^\ell - y\|^2_{L^2(0,T;X)}}{\|y^{\ell+1} - y\|^2_{L^2(0,T;X)}} \Rightarrow EOD := \frac{1}{\ell_{\max}} \sum_{k=1}^{\ell_{\max}} Q(k) \approx \alpha.$$

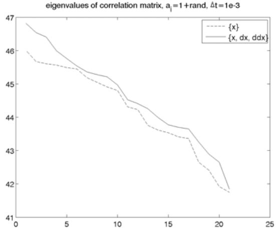
Decay of eigenvalues and of norms

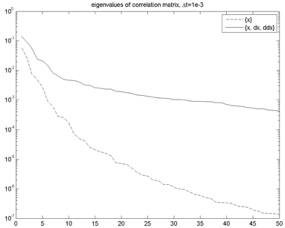




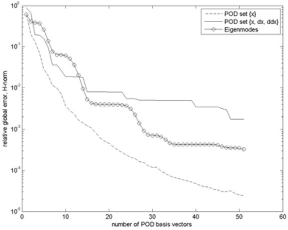
POD for wave equation - decay of modes and error

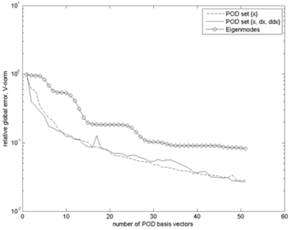
Decay of eigenvalues; without damping (left), and with damping (right)





Errors: H-modes (left) and V-modes (right), compared to Fourier analysis.





Shortcomings of POD - non-smooth systems

The Cahn-Hilliard system

$$\begin{aligned} \partial_t \varphi - m\Delta\mu + v \cdot \nabla\varphi &= 0, \\ -\sigma\varepsilon\Delta\varphi + \sigma\varepsilon^{-1}\mathcal{F}'(\varphi) &= \mu. \end{aligned} \tag{CH}$$

weak form:

$$\underbrace{\begin{aligned} \langle \partial_t \varphi, \Phi\rangle + \langle v \cdot \nabla\varphi, \Phi\rangle + m\langle \nabla\mu, \nabla\Phi\rangle &= 0 \\ -\langle \mu, \Psi\rangle + \sigma\varepsilon\langle \nabla\varphi, \nabla\Psi\rangle + \frac{\sigma}{\varepsilon}\langle \mathcal{F}'(\varphi), \Psi\rangle &= 0 \end{aligned}}_{=:\langle F(\varphi,\mu),(\Phi,\Psi)\rangle}$$

relaxed Double Obstacle Energy:

$$\mathcal{F}(\varphi) = \frac{1}{2}\left(1 - \varphi^2\right) + \frac{s}{k}\left(\max\left(\varphi - 1, 0\right) + |\min\left(\varphi + 1, 0\right)|\right)^k \quad k \in \mathbb{N}$$

Decay of modes depends on the smoothness of the potential

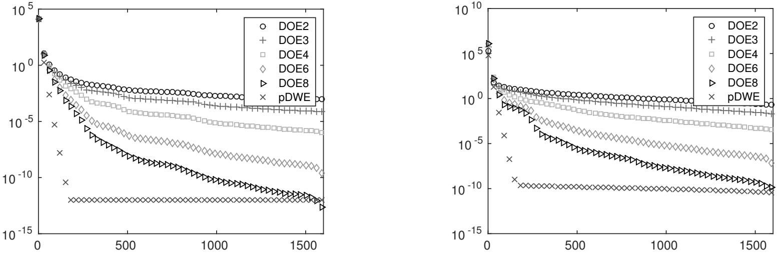


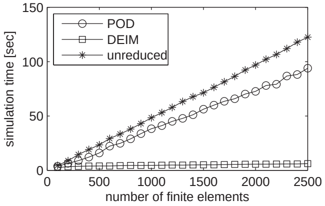
Figure: Singular values: ϕ (left), μ (right)

Nonlinearities - DEIM by Chaturantabut and Sorensen (SISC 2010)

POD projects the nonlinearity $\mathcal{G}(y)$ in the PDE as follows:

$$\mathcal{G}^{\ell}(\alpha(t)) \equiv \underbrace{\Phi^t}_{\ell \times N} \underbrace{\mathcal{G}(\Phi\alpha(t))}_{N \times 1}.$$

Here, Φ is $N \times \ell$, with N the dimension of the finite element space, $\mathcal{G}$ has N components, and in the evaluation of every of its components may touch every component of its $N-$dimensional argument. This evaluation thus has complexity $\mathcal{O}(\ell N)$.



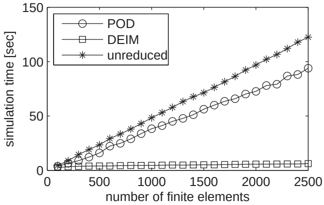
POD versus POD-DEIM in MOR for semiconductors governed by the Drift-Diffusion model

Nonlinearities - DEIM by Chaturantabut and Sorensen (SISC 2010)

POD projects the nonlinearity $\mathcal{G}(y)$ in the PDE as follows:

$$\mathcal{G}^\ell(\alpha(t)) \equiv \underbrace{\Phi^t}_{\ell\times N} \underbrace{\mathcal{G}(\Phi\alpha(t))}_{N\times 1}.$$

Here, Φ is $N \times \ell$, with N the dimension of the finite element space, $\mathcal{G}$ has N components, and in the evaluation of every of its components may touch every component of its $N-$dimensional argument. This evaluation thus has complexity $\mathcal{O}(\ell N)$.



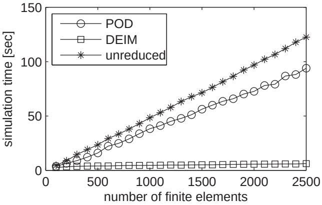
POD versus POD-DEIM in MOR for semiconductors governed by the Drift-Diffusion model

Nonlinearities - DEIM by Chaturantabut and Sorensen (SISC 2010)

POD projects the nonlinearity $\mathcal{G}(y)$ in the PDE as follows:

$$\mathcal{G}^{\ell}(\alpha(t)) \equiv \underbrace{\Phi^t}_{\ell \times N} \underbrace{\mathcal{G}(\Phi\alpha(t))}_{N \times 1}.$$

Here, Φ is $N \times \ell$, with N the dimension of the finite element space, $\mathcal{G}$ has N components, and in the evaluation of every of its components may touch every component of its $N-$dimensional argument. This evaluation thus has complexity $\mathcal{O}(\ell N)$.



POD versus POD-DEIM in MOR for semiconductors governed by the Drift-Diffusion model

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.

Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \ldots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \ldots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \ldots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \ldots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m} \underbrace{P^t\mathcal{G}}_{m \text{ evals}} \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \le \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\rightarrow$ improve error bound through modifying P, and thus DEIM.

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.
Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \ldots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \ldots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \ldots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \ldots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m} \underbrace{P^t\mathcal{G}}_{m \text{ evals}} \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \leq \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\rightarrow$ improve error bound through modifying P, and thus DEIM.

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.
Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \dots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \dots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \dots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \dots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m} \underbrace{P^t\mathcal{G}}_{m \text{ evals}} \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \le \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\rightarrow$ improve error bound through modifying P, and thus DEIM.

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.
Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \ldots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \ldots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \ldots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \ldots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m} \underbrace{P^t\mathcal{G}}_{m \text{ evals}} \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \leq \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\rightarrow$ improve error bound through modifying P, and thus DEIM.

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.
Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \ldots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \ldots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \ldots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \ldots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m}\ \underbrace{P^t\mathcal{G}}_{m \text{ evals}}\ \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \leq \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\to$ improve error bound through modifying P, and thus DEIM.

DEIM-idea

Approximate the nonlinear function $\mathcal{G}(\Phi\alpha(t))$ by projecting it onto a subspace that approximates the space generated by the nonlinear function and that is spanned by a basis of dimension $m << N$.
Here: perform a SVD with $Y := [\mathcal{G}(y(t_1)), \ldots, \mathcal{G}(y(t_n))]$ and use the first m modes $U := [u_1, \ldots, u_m]$ to interpolate

$$\mathcal{G}(\Phi\alpha(t)) \approx Uc(t).$$

This system is overdetermined.
Now DEIM selects m rows $\rho_1, \ldots, \rho_m$ from this system by a *greedy procedure*;

$$P^t\mathcal{G}(\Phi\alpha(t)) \approx (P^tU)c(t), \text{ where } P := [e_{\rho_1}, \ldots, e_{\rho_m}] \in \mathbb{R}^{N\times m},$$

with P^tU invertible, so that $c(t)$ is uniquely determined.
This gives

$$\mathcal{G}^\ell(\alpha(t)) \approx \underbrace{\Phi^tU(P^tU)^{-1}}_{\ell\times m}\ \underbrace{P^t\mathcal{G}}_{m \text{ evals}}\ \underbrace{(\Phi\alpha(t))}_{N\times\ell} =: \hat{\mathcal{G}}^\ell(\alpha(t))$$

with the error bound

$$\|\mathcal{G}^\ell - \hat{\mathcal{G}}^\ell\|_2 \leq \|(P^tU)^{-1}\|_2\|(I - UU^t)\mathcal{G}^\ell\|_2.$$

Serkan's talk $\rightarrow$ improve error bound through modifying P, and thus DEIM.

Upcoming: Optimization with POD surrogate models

Upcoming: Optimization with POD surrogate models

Thank you for attending!

Lecture 2: Optimization with POD surrogate models

Lecture 2: Optimization with POD surrogate models

Optimization with the reduced model - the beginnings

Model optimization problem:

$$\min_{(y,u)\in W\times U} J(y,u) := \frac{1}{2}\int_{Q_o} |y-z|^2 \, dxdt + \frac{\gamma}{2}\|u\|_U^2$$

s.t.

$$\begin{aligned}
\frac{\partial y}{\partial t} + (y\cdot\nabla)y - \nu\Delta y + \nabla p &= Bu \text{ in } Q = (0,T)\times\Omega,\\
-\text{div } y &= 0 \text{ in } Q,\\
y(t,\cdot) &= 0 \text{ on } \Sigma = (0,T)\times\partial\Omega,\\
y(0,\cdot) &= y_0 \text{ in } \Omega.
\end{aligned}$$

Here, $B : U \to L^2(0,T;H^{-1}(\Omega)^d)$ denotes the control operator. It is also possible to consider the initial values as control.

Typical control operator is extension $B : L^2(0,T;L^2(\Omega_c)^d) \to L^2(0,T;H^{-1}(\Omega)^d)$. Observation cylinder is given by $Q_o := (0,T)\times\Omega_o$.

POD model as pde surrogate in the optimization problem

Ansatz for state (and the desired state)

$$y = \bar{y} + \sum_{i=1}^{n} \alpha_i \Phi_i, \quad z = \bar{y} + \sum_{i=1}^{n} \alpha_i^z \Phi_i.$$

Optimization problem with POD surrogate model

$$\min_{(y,u)} J(y,u) = J(\alpha,u) = \frac{1}{2}\int_0^T (\alpha - \alpha^z)^t \mathcal{M}_1 (\alpha - \alpha^z)\, dt + \frac{\gamma}{2}\|u\|_U^2$$

s.t.

$$\begin{aligned} \dot{\alpha} + A\alpha + n(\alpha) &= r + \mathcal{B}u, \\ \alpha(0) &= a_0. \end{aligned}$$

Validity of surrogate model

Fact:
Control changes system dynamics.

Consequence:
Mean and modes should be suitably modified during the optimization process.

Idea:
Adaptively modify the surrogate model and thus, the reduced optimization problems.

Adaptive POD control – Afanasiev, Hinze 1999

1. **Snapshots $y_i^0, i = 1, \ldots, N_0$ given, u^0 given control, $\delta \in [0, 1]$ required relative information content, j=0.**

2. **Compute $M = \operatorname{argmin} \left\{ \mathrm{I}(M) := \sum_{k=1}^{M} \lambda_k \,/\, \sum_{k=1}^{N} \lambda_k;\ \mathrm{I}(M) \geq \delta \right\}$.**

3. **Compute POD modes and solve**

$$
(\mathrm{ROM}) \begin{cases} \min J(\alpha, u) \\ \text{s.t.} \\ \dot{\alpha} + A\alpha + n(\alpha) = \mathcal{B}u, \quad \alpha(0) = a_0. \end{cases}
$$

   for u^j.

4. **Compute y^j corresponding to Bu^j and new snapshots $y_i^{j+1}, i = N_j + 1, \ldots, N_{j+1}$ to the snapshot set $y_i^j, i = 1, \ldots, N_j$.**

5. **While $\|u^{j+1} - u^j\|_U$ is large, j = j+1 and goto 2.**

Numerical comparison

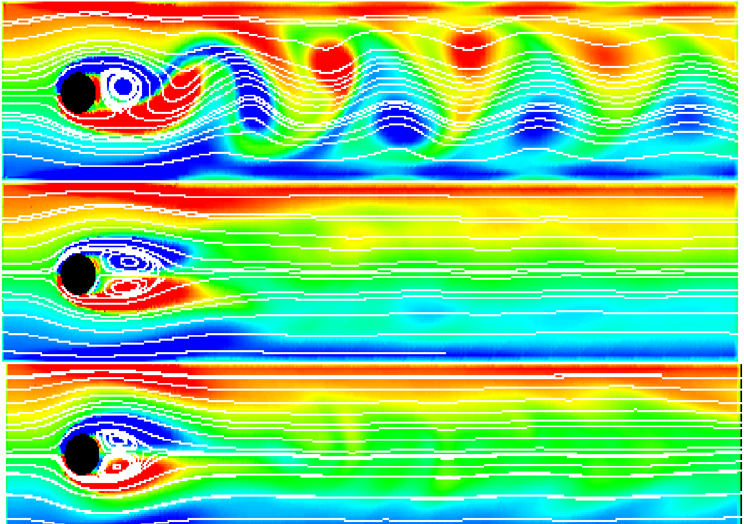
Flow around a circular cylinder at Re=100. Control gain: Tracking of Stokes flow (or mean flow) $\bar{y}$ in an observation volume Ω_{obs} behind the cylinder by applying a volume force in the control volume Ω_c.
Cost functional:

$$J(y,u) = \frac{\gamma}{2} \int_0^T \int_{\Omega_c} |u|^2 \, dxdt + \frac{1}{2} \int_0^T \int_{\Omega_{obs}} |y - \bar{y}|^2 \, dxdt$$

CPU time needed to compute the suboptimal controls $\approx$ 40 times smaller than that needed to compute the optimal open loop control. But the quality of the controls is very similar.

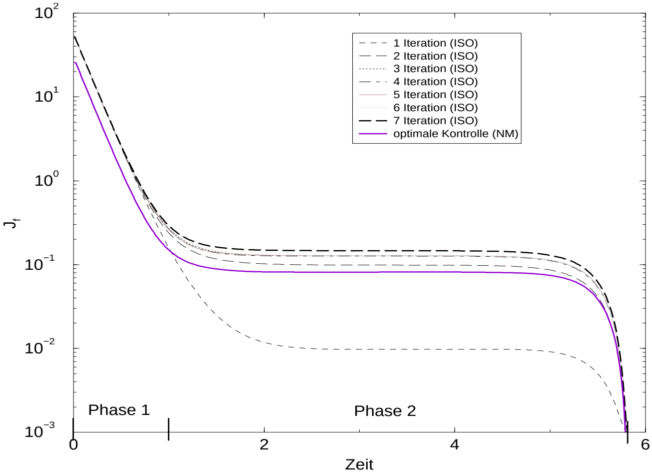
$$\text{Runtime(Optimization Problem)} = 6 - 8 \times \text{Runtime(PDE)}$$

Uncontrolled flow, target flow = mean flow, controlled flow at $t = 3.4$.



Numerical results cont.

Control cost, $\Omega_o = \Omega = \Omega_c$, tracking of mean flow



Motivation: optimization problem with pde constraints

$$\min_{(y,u)\in W\times U_{ad}} J(y,u) \text{ s.t.}$$

$$\begin{aligned} \frac{\partial y}{\partial t} + \mathcal{A}y + \mathcal{G}(y) &= Bu \text{ in } Z^* \\ y(0) &= y_0 \text{ in } H. \end{aligned}$$

Approach: Solve this problem by using a POD surrogate model;

$$\min_{(y^l,u^l)\in W^l\times U_{ad}} J^l(y^l,u^l) \text{ s.t.}$$

$$\begin{aligned} \frac{\partial y^l}{\partial t} + \mathcal{A}^l y^l + \mathcal{G}^l(y^l) &= Bu^l \text{ in } (Z^l)^* \\ y^l(0) &= y_0^l \text{ in } H^l. \end{aligned}$$

Tasks:

- **Error estimation,**
- **adaption of the POD surrogate model during the optimization loop.**

Mathematical setting, state equation

- V, H separable Hilbert spaces, $(V, H = H^*, V^*)$ Gelfand triple.
- $a : V \times V \to \mathbb{R}$ bounded, coercive and symmetric. Set
$$\langle \bullet, \bullet \rangle_V := a(\bullet, \bullet).$$
- U Hilbert space, $B : U \to \mathcal{L}^2(U, L^2(V^*))$ linear control operator, $y_0 \in H$.
- State equation
$$\begin{aligned} \frac{d}{dt}(y(t), v)_H + a(y(t), v) &= \langle (Bu)(t), v \rangle_{V,V^*}, \quad t \in [0, T], v \in V, \\ (y(0), v)_H &= (y_0, v)_H, \quad v \in V. \end{aligned}$$
- For every $u \in U$ the solution $y = y(u) \in W := \{w \in L^2(V), w_t \in L^2(V^*)\}$ is unique.

Optimization problem

- **Cost functional**

$$J(y,u) := \frac{1}{2}\|y - z\|^2_{L^2(H)} + \frac{\alpha}{2}\|u\|^2_U.$$

- **Admissibility:** $u \in U_{\text{ad}} \subseteq U$ closed, convex, $y \equiv y(u)$ unique solution of state equation associated to u, i.e.

$$\begin{aligned} \frac{d}{dt}(y(t),v)_H + a(y(t),v) &= \langle (Bu)(t), v\rangle_{V,V^*}, & t \in [0,T], v \in V, \\ (y(0),v)_H &= (y_0, v)_H, & v \in V. \end{aligned}$$

- **Minimization problem:**

$$(P) \quad \min_{(y,u)\in W(0,T)\times U_{\text{ad}}} J(y,u) \text{ s.t. Admissibility.}$$

- (P) admits a unique solution $(y,u) \in W \times U_{\text{ad}}$.

Optimality conditions

- With the reduced cost functional $\hat{J}(u) := J(y(u), u)$ there holds

$$\left(\hat{J}'(u), v - u\right) \geq 0 \text{ for all } v \in U_{\text{ad}}.$$

- Here

$$\hat{J}'(u) = \alpha u + B^* p(y(u)).$$

- The function p solves the adjoint equation

$$\begin{aligned} -\tfrac{d}{dt}(p(t), v)_H + a(v, p(t)) &= (y - z, v)_H, & t \in [0, T], v \in V, \\ (p(T), v)_H &= 0, & v \in V. \end{aligned}$$

- Variational inequality equivalent to nonsmooth operator equation

$$u = P_{U_{\text{ad}}}\left(-\frac{1}{\alpha} B^* p(y(u))\right)$$

with $P_{U_{\text{ad}}}$ denoting the orthogonal projection onto U_{ad}.

Discrete concept for the state equation

- For $l \in \mathbb{N}$ choose a POD subspace $V^l := \langle \chi_1, \ldots, \chi_l \rangle$ of V with the property

$$\|y(t) - \sum_{k=1}^{l} (y(t), \chi_k)_V \, \chi_k\|^2_{W(0,T)} \sim \sum_{k=l+1}^{\infty} \lambda_k .$$

- Galerkin semi-discretization y^l of state y using subspace V^l:

$$\begin{aligned} \tfrac{d}{dt}\left(y^l(t), v\right)_H + a(y^l(t), v) &= \langle (Bu)(t), v \rangle_{V,V^*}, & t \in [0,T], v \in V^l, \\ (y(0), v)_H &= (y_0, v)_H , & v \in V^l. \end{aligned}$$

- If needed, define similarly a Galerkin semi-discretization p^l of p;

$$\begin{aligned} -\tfrac{d}{dt}\left(p^l(t), v\right)_H + a(v, p^l(t)) &= \left(y^l - z, v\right)_H , & t \in [0,T], v \in V^l, \\ \left(p^l(T), v\right)_H &= 0, & v \in V^l. \end{aligned}$$

Optimization problem with POD surrogate model

- **Discrete minimization problem:**

$$(\hat{P}^l) \qquad \min_{u \in U_{\mathrm{ad}}} \hat{J}^l(u) := J(y^l(u), u).$$

- $(\hat{P}^l)$ **admits a unique solution** $u^l \in U_{\mathrm{ad}}$.
- **Optimality condition:**

$$\left(\hat{J}^{l\prime}(u), v - u^l\right) \geq 0 \text{ for all } v \in U_{\mathrm{ad}}.$$

- **Here**

$$\hat{J}^{l\prime}(u) = \alpha u + B^* p^l(y^l(u)).$$

- **The function** p^l **solves the adjoint equation**

$$\begin{aligned} -\tfrac{d}{dt}\left(p^l(t), v\right)_H + a(v, p^l(t)) &= \left(y^l - z, v\right)_H, \quad && t \in [0, T], v \in V^l, \\ \left(p^l(T), v\right)_H &= 0, && v \in V^l. \end{aligned}$$

- **Variational inequality equivalent to nonsmooth operator equation**

$$u^l = P_{U_{\mathrm{ad}}}\left(-\frac{1}{\alpha} B^* p^l(y^l(u))\right).$$

Error estimate

Theorem: Let u, u^l denote the unique solutions of (P) and $(\hat{P}^l)$, respectively. Then

$$\|u - u^l\|_U^2 \leq \frac{1}{\alpha}\Big\{\Big(B^*(p(y(u)) - p^l(y(u))), u^l - u\Big)_U + \\ + \int_0^T \Big(y^l(u^l) - y^l(u), y(u) - y^l(u)\Big)_H dt\Big\}$$

Using the analysis of Kunisch and Volkwein for POD approximations one gets

$$\|u - u^l\|_U \sim \|y_0 - P^l y_0\|_H + \sqrt{\sum_{k=l+1}^{\infty} \lambda_k} + \\ + \|y_t - \mathcal{P}^\ell y_t\|_{L^2(0,T;V')} + \|p(y(u)) - P^l(p(y(u)))\|_{W(0,T)}$$

Conclusions from the analysis

- Get rid of $\|(y-\mathcal{P}^\ell y)_t\|^2_{L^2(0,T;V')}$ $\rightarrow$ include derivative information into your snapshot set.
- Get rid of $\|p-\mathcal{P}^\ell p\|^2_{W(0,T)}$ $\rightarrow$ include adjoint information into your snapshot set.

Recipe:

For $l \in \mathbb{N}$ choose a POD subspace $V^l := \langle \chi_1, \ldots, \chi_l \rangle$ of V with the property

$$\|y(t) - \sum_{k=1}^{l} (y(t), \chi_k)_V \, \chi_k\|^2_{W(0,T)} \sim \sum_{k=l+1}^{\infty} \lambda_k,$$

and if one intends to solve optimization problems, also ensure

$$\|p(t) - \sum_{k=1}^{l} (p(t), \chi_k)_V \, \chi_k\|^2_{W(0,T)} \sim \sum_{k=l+1}^{\infty} \lambda_k,$$

Conclusions from the analysis

- **Get rid of $\|(y-\mathcal{P}^\ell y)_t\|^2_{L^2(0,T;V')}$ → include derivative information into your snapshot set.**
- **Get rid of $\|p-\mathcal{P}^\ell p\|^2_{W(0,T)}$ → include adjoint information into your snapshot set.**

Recipe:

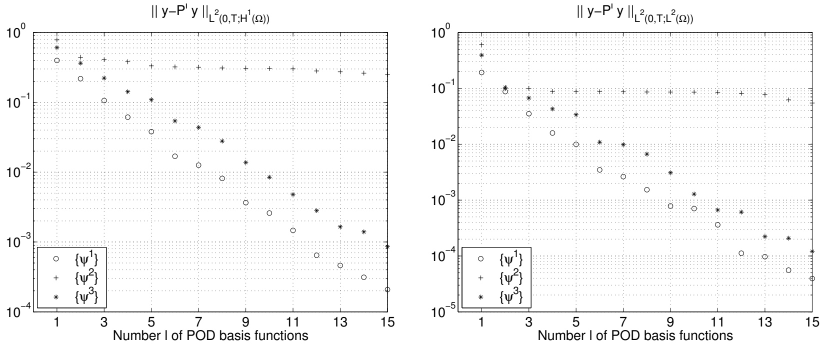
For $l \in \mathbb{N}$ choose a POD subspace $V^l := \langle \chi_1, \dots, \chi_l \rangle$ of V with the property

$$\|y(t) - \sum_{k=1}^{l} (y(t), \chi_k)_V \, \chi_k\|^2_{W(0,T)} \sim \sum_{k=l+1}^{\infty} \lambda_k,$$

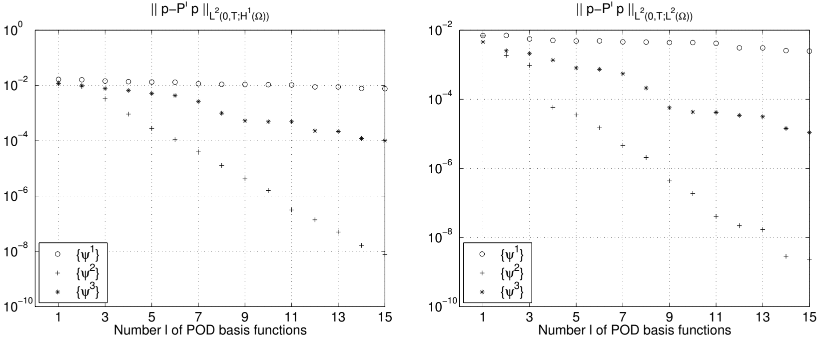
and if one intends to solve optimization problems, also ensure

$$\|p(t) - \sum_{k=1}^{l} (p(t), \chi_k)_V \, \chi_k\|^2_{W(0,T)} \sim \sum_{k=l+1}^{\infty} \lambda_k,$$

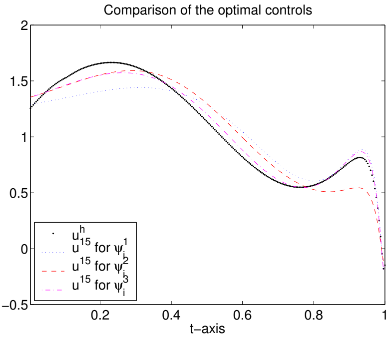
Error between state and and its orthogonal projection

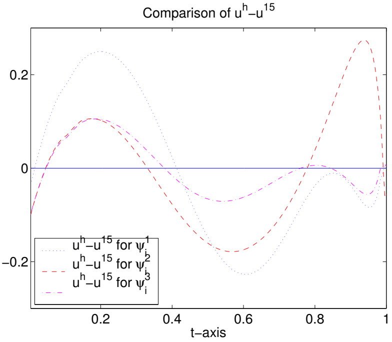


Error between co-state and its orthogonal projection



Neumann boundary control of the heat equation





Snapshot location for parabolic (*mother*) optimal control problem

We consider optimal distributed control of the linear heat equation. If the data of the optimal control problem is smooth enough we have $\alpha u + p = 0$ and

1. **the optimal state y satisfies**

$$\begin{aligned} -\frac{\partial^2 y}{\partial t^2} + \Delta^2 y + \frac{1}{\alpha} y = \frac{1}{\alpha} z \quad &\text{in } \Omega_T, \\ y = 0 \quad &\text{on } \Sigma_T, \\ \Delta y = 0 \quad &\text{on } \Sigma_T, \\ (\frac{\partial y}{\partial t} - \Delta y)(T) = 0 \quad &\text{in } \Omega, \\ y(0) = y_0 \quad &\text{in } \Omega, \end{aligned}$$

2. **while p solves**

$$\begin{aligned} -\frac{\partial^2 p}{\partial t^2} + \Delta^2 p + \frac{1}{\alpha} p = -\frac{\partial z}{\partial t} + \Delta z \quad &\text{in } \Omega_T, \\ p = 0 \quad &\text{on } \Sigma_T, \\ \Delta p = z \quad &\text{on } \Sigma_T, \\ (\frac{\partial p}{\partial t} + \Delta p)(0) = y_d(0) - y_0 \quad &\text{in } \Omega, \\ p(T) = 0 \quad &\text{in } \Omega. \end{aligned}$$

Snapshot location in parabolic optimal control

With y, p and y_k, p_k time-discrete approximations to y, p we have

$$\|y - y_k\|^2_{2,1,\Omega_T} \leq C\eta_y^2,$$

where

$$\eta_y^2 = \sum_n k_n^2 \int_{I_n} \|\frac{1}{\alpha}y_d + \frac{\partial^2 y_k}{\partial t^2} - \frac{1}{\alpha}y_k - \Delta^2 y_k\|^2_{0,\Omega} + \sum_n \int_{I_n} \|\Delta y_k\|^2_{0,\Gamma},$$

and

$$\|p - p_k\|^2_{2,1,\Omega_T} \leq C\eta_p^2,$$

where

$$\eta_p^2 = \sum_n k_n^2 \int_{I_n} \|-\frac{\partial y_d}{\partial t} + \Delta y_d + \frac{\partial^2 p_k}{\partial t^2} - \frac{1}{\alpha}p_k - \Delta^2 p_k\|^2_{0,\Omega} + \sum_n \int_{I_n} \|y_d - \Delta p_k\|^2_{0,\Gamma}.$$

Idea: go for an adaptive time grid, based on a coarse discretization in space, and use this time-grid as snapshot grid for the optimal control problem.

Snapshot location in parabolic optimal control - numerical example

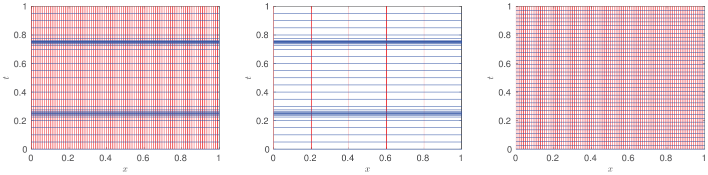


FIGURE 8. Test 6.2: Space-time grid with dof = 37 according to the strategy in [5] with $\Delta x = 1/100$ (left) and $\Delta x = 1/5$ (middle), respectively, and equidistant grid (right)

Δt	$\varepsilon^y_{\text{abs}}$	$\varepsilon^u_{\text{abs}}$	dof	$\varepsilon^y_{\text{abs}}$	$\varepsilon^u_{\text{abs}}$
1/24	$8.5101 \cdot 10^{+00}$	$6.1244 \cdot 10^{-01}$	25	$2.5927 \cdot 10^{+00}$	$2.0417 \cdot 10^{-01}$
1/36	$5.5089 \cdot 10^{+00}$	$3.9009 \cdot 10^{-01}$	37	$4.1726 \cdot 10^{-01}$	$2.9296 \cdot 10^{-02}$
1/76	$2.2935 \cdot 10^{+00}$	$1.5923 \cdot 10^{-01}$	77	$1.9320 \cdot 10^{-01}$	$2.0327 \cdot 10^{-02}$
1/148	$1.1526 \cdot 10^{+00}$	$8.0239 \cdot 10^{-02}$	149	$4.8640 \cdot 10^{-02}$	$1.9035 \cdot 10^{-02}$

TABLE 2. Test 6.2: Absolute errors between the exact optimal solution and the POD suboptimal solution depending on the time discretization (equidistant: columns 1-3, adaptive: columns 4-6)

Recent developments–TRPOD by Arian, Fahl and Sachs 2000–

Idea: Use a POD surrogate model as model function in the Trust–Region process. Let

$$J(u) = J(y(u), u), \quad \hat{J}(u) = J(\hat{y}(u), u),$$

with $\hat{y}(u)$ the response of the POD surrogate model.
Pseudo Algorithm:

1. Given u, compute POD model
2. Compute $s^* = \operatorname{argmin}_{\|u-s\| \leq \Delta} \hat{J}(u+s)$
3. $$\rho := \frac{J(u+s^*) - J(u)}{\hat{J}(u+s^*) - \hat{J}(u)} \quad \begin{cases} \text{large:} & u = u + s^*, \quad \text{increase } \Delta \\ \text{moderate:} & u = u + s^*, \quad \text{decrease } \Delta \\ \text{small:} & \text{keep } u, \quad \text{decrease } \Delta \end{cases}$$

Global convergence under standard TR assumptions plus $\frac{\|J'(u) - \hat{J}'(u)\|}{\|\hat{J}'(u)\|}$ sufficiently small.

Recent developments–OSPOD by Kunisch and Volkwein 2006

Idea: Include choice of trajectory dependent POD modes as subsidiary condition into the optimization problem. This reads

$$(P^l_{OSPOD}) \quad \begin{cases} \min_{\alpha,\Phi,u} \hat{J}(\alpha,\Phi,u) \text{ s.t.} \\ M(\Phi)\dot{\alpha} + A(\Phi)\alpha + n(\Phi)(\alpha) = B(\Phi)u, \\ M(\Phi)\alpha(0) = \alpha_0(\Phi), \\ y_t + \mathcal{A}y + \mathcal{G}(y) = \mathcal{B}u, \\ y(0) = y_0, \\ \mathcal{R}(y)\Phi_i = \lambda_i\Phi_i \text{ for } i = 1,\dots,l, \\ \|\Phi_i\|_X = 1 \text{ for } i = 1,\dots,l. \end{cases}$$

Here $\Phi = [\Phi_1,\dots,\Phi_l]$, $y^l = \sum_{i=1}^{l} \alpha_i(t)\Phi$, **and**

$$\mathcal{R}(y)(z) := \int_0^T \langle y(t), z\rangle_X y(t)dt \text{ for } z \in X.$$

A very similar approach is proposed by Ghattas, van Bloemen Waanders and Willcox 2005.

How many snapshots?

Meyer, Matthies, Heuveline, H.

How many snapshots? $\longrightarrow$ iterative goal oriented procedure.

- **Goal: Resolve J(y)**
- **Start on coarse equi-distant time grid and compute snapshots**
- **Build POD model and compute y_h and adjoint z_h of reduced dynamics**
- **Becker and Rannacher: $J(y) - J(y_h) \approx \eta(y_h, z_h)$**
- **$\eta(y_h, z_h) >$ tol: double number of snapshots (re-computation)**

Where to take snapshots?

Where to take snapshots? $\longrightarrow$ time-step adaption via sensitivity of POD model.

- **Goal: Optimal time-grid for system dynamics**
- **Start on coarse (equi-distant) time grid and compute snapshots**
- **Build POD model and compute y_h and adjoint z_h of reduced dynamics**
- **Becker, Johnson, Rannacher:** $\eta(y_h, z_h) = \sum_{I_j} \rho_j^{loc}(y_h)\omega_j^{loc}(z_h)$
- **New time-grid: equi-distribute** $\rho_j^{loc}(y_h)\omega_j^{loc}(z_h)$

Further developments and improvements

- Efficient treatment of nonlinearities → Chaturantabut, Sorensen (2010)
- MOR for the input–output map → Heiland, Mehrmann
- A posteriori POD concept → Tröltzsch and Volkwein (2010)
- Which modes? → DWR concepts (Matthies, Meyer 2003)
- How many snapshots? → iterative goal oriented DWR procedure
- Were take snapshots? → time–step adaption via sensitivity of the POD model
- POD in the context of space–mapping
- Sampling of parameter (≡ control) space → Greedy sampling by Patera and Rozza (2007)
- Use of *linear* MOR techniques for nonlinear problems → SQP context, semi–linear time integration, domain decomposition

Thank you for your attention

Further developments and improvements

- Efficient treatment of nonlinearities → Chaturantabut, Sorensen (2010)
- MOR for the input–output map → Heiland, Mehrmann
- A posteriori POD concept → Tröltzsch and Volkwein (2010)
- Which modes? → DWR concepts (Matthies, Meyer 2003)
- How many snapshots? → iterative goal oriented DWR procedure
- Were take snapshots? → time–step adaption via sensitivity of the POD model
- POD in the context of space–mapping
- Sampling of parameter ($\equiv$ control) space → Greedy sampling by Patera and Rozza (2007)
- Use of *linear* MOR techniques for nonlinear problems → SQP context, semi–linear time integration, domain decomposition

Thank you for your attention