The Application of Proper Orthogonal Decomposition to the Control of Tubular Reactors

Doctoral Presentation

Oscar Mauricio Agudelo Mañozca

Jury :

Y. Willems, chairman

B. De Moor, promotor

J. Espinosa, co-promotor (UNAL)

J. Vandewalle, co-promotor

J. Suykens

J. Van Impe

A.C.P.M. Backx (TU/e)

I. Smets

SCD Research Division
ESAT- K. U. Leuven
November 10th, 2009

K. U. LEUVEN

ESAT

Outline

- PART I: Application of Proper Orthogonal Decomposition (POD) in the design of predictive controllers for tubular chemical reactors
- PART II: Acceleration of the evaluation of nonlinear POD models
- Conclusions and Future Research

PART I

Application of Proper Orthogonal Decomposition (POD) in the design of predictive controllers for tubular chemical reactors

Motivation

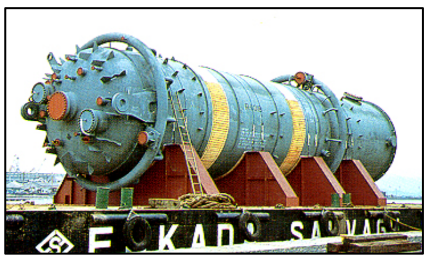
- Current control schemes for Tubular Reactors:
  - Overly conservative.
  - Do not push the plants to their limits of performance.

- In this research:

  "We develop advanced POD-based predictive control strategies".

Proposed control design framework:

Tubular Reactors



Modeled by

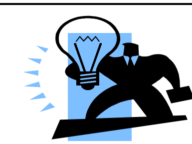
Partial Differential Equations

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial z} - k_0 C e^{-\frac{E}{RT}}$$

$$\frac{\partial T}{\partial t} = -v\frac{\partial T}{\partial z} + G_r C e^{-\frac{E}{RT}} + H_r (T_w - T)$$

Discretization

High Order Models
1000, 10.000 states, etc

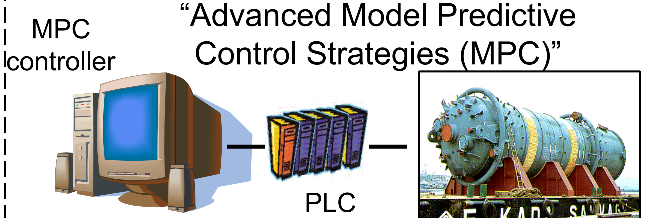


Proper Orthogonal Decomposition and Galerkin Projection

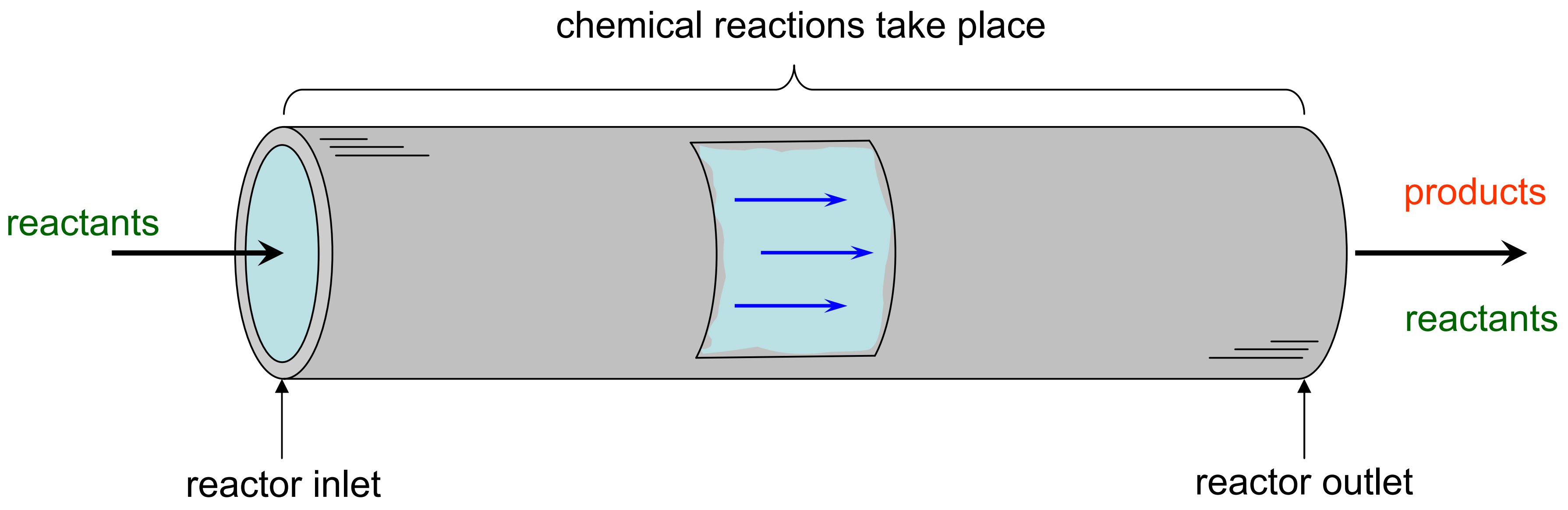
Reduced-order Models

Allow the design of

"Advanced Model Predictive Control Strategies (MPC)"



What is a Tubular Chemical Reactor ?

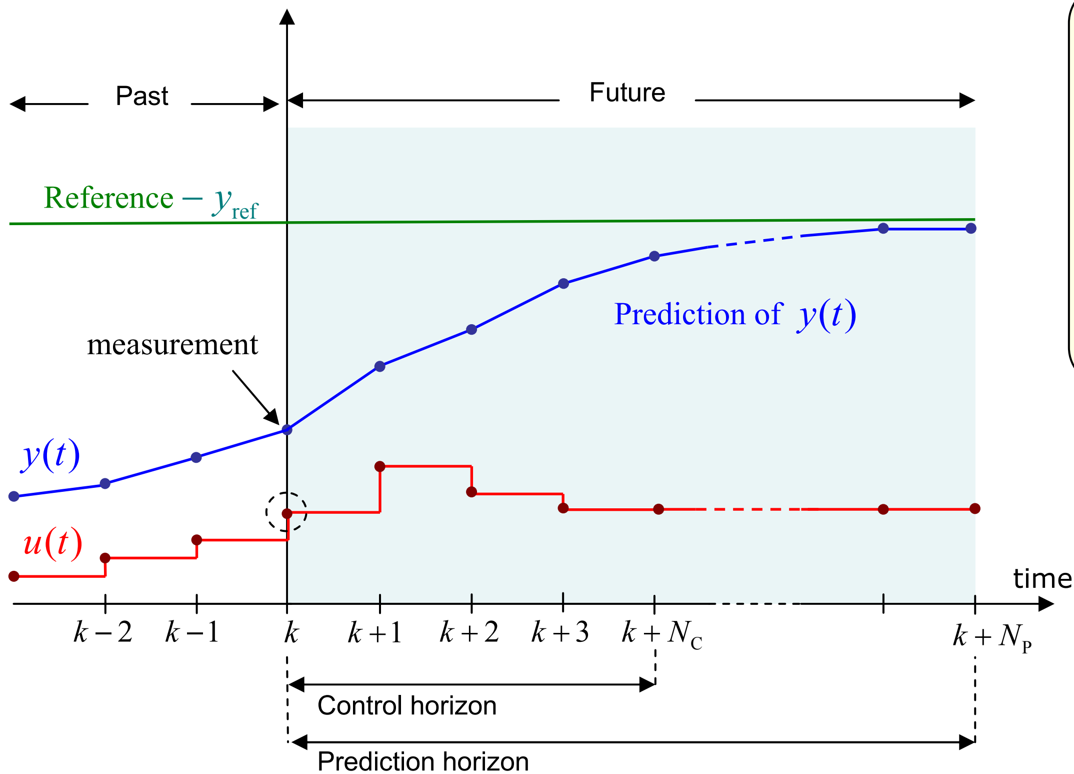


- They typically operate in steady-state regimes ➔ to produce high product volumes of a consistent quality.
- They pose interesting control problems :
  - Their behavior is modeled by highly nonlinear Partial Differential Equations (PDEs)
  - Satisfaction of Input and State constraints.

Model Predictive Control (MPC)

Control method for handling input and state constraints within an optimal control setting.

Principle of predictive control



$$\min_{u(k),\dots,u(k+N_c-1)} \sum_{i=1}^{N_p} \left(y_{\text{ref}} - y(k+i)\right)^2$$

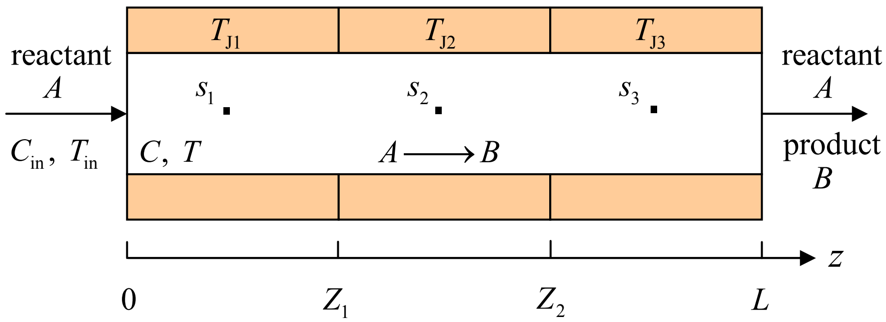
subject to

- model of the process
- input constraints
- output / state constraints

Why MPC ?

- It handles multivariable interactions
- It handles input and state constraints
- It can push the plants to their limits of performance.

Process: Non-isothermal tubular reactor



The system to be controlled is a non-isothermal tubular reactor where a single, first order, irreversible, exothermic reaction takes place ($A \rightarrow B$). A Plug-flow behavior is assumed.

The mathematical model is given by the following coupled-nonlinear PDEs :

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial z} - k_0 C e^{-\frac{E}{RT}}$$

$$\frac{\partial T}{\partial t} = -v\frac{\partial T}{\partial z} + G_r C e^{-\frac{E}{RT}} + H_r(T_w - T)$$

$$G_r = -\frac{\Delta H k_0}{\rho C_p}, \quad H_r = \frac{4h}{d\rho C_p}$$

$$T_w = \begin{cases} T_{J1}, & 0 \le z < Z_1 \\ T_{J2}, & Z_1 \le z < Z_2 \\ T_{J3}, & Z_2 \le z \le L \end{cases}$$

$$C = C_{in} \text{ at } z = 0$$

$$T = T_{in} \text{ at } z = 0$$

Input Constraints: $280\text{ K} \le T_{J1}, T_{J2}, T_{J3} \le 400\text{ K}$

State Constraints: $T(z,t) \le 400\text{ K}$

Disturbances in the feed flow :

$T_{in}^{\Delta} = \pm 10\text{K}$ for T_{in}

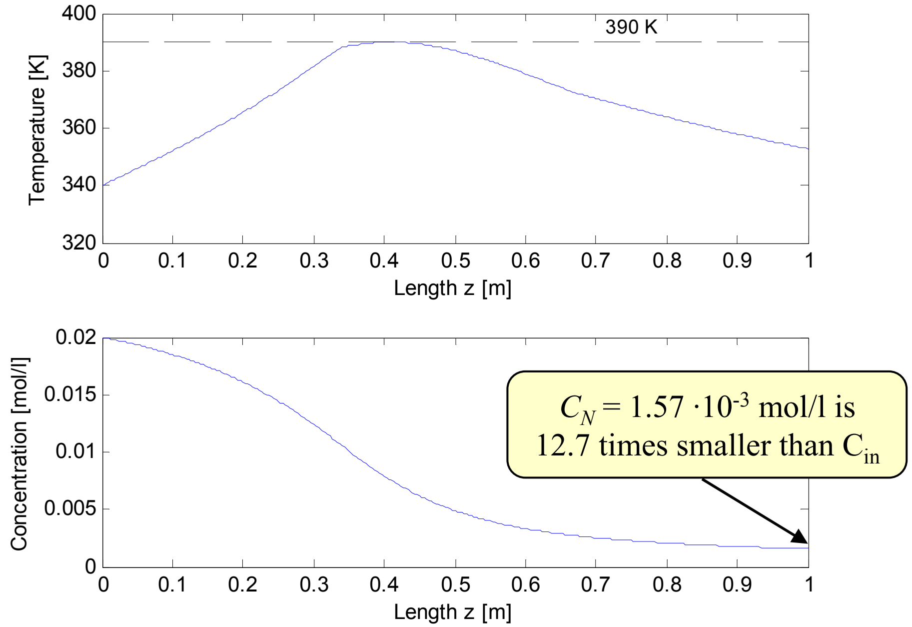
$C_{in}^{\Delta} = \pm$ 5% of the nominal value of C_{in}

$T(z,t)$ = reactant temperature in [K], $C(z,t)$ = reactant concentration in [mol/l]

Process: Non-isothermal tubular reactor

Operating profiles of the reactor

(T_{J1}= 374.6 K, T_{J2}=310.1 K and T_{J3}=352.2 K)



The discretized (in space) linear model of the reactor around the operating profiles is given by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{F}\mathbf{d}(t)$$

where :

$$\mathbf{x}(t) = \left[\bar{C}_1^{\Delta}, \bar{C}_2^{\Delta}, \ldots, \bar{C}_N^{\Delta}, \bar{T}_1^{\Delta}, \bar{T}_2^{\Delta}, \ldots, \bar{T}_N^{\Delta}\right]^T$$

$$\mathbf{d}(t) = \left[\bar{C}_{\text{in}}^{\Delta}, \bar{T}_{\text{in}}^{\Delta}\right]^T$$

$$\mathbf{u}(t) = \left[\bar{T}_{\text{J1}}^{\Delta}, \bar{T}_{\text{J2}}^{\Delta}, \bar{T}_{\text{J3}}^{\Delta}\right]^T$$

$\Delta \rightarrow$ deviation from steady state. $\bar{\ } \rightarrow$ Normalized.

The **control goal** is to keep the reactor around a desired operating condition in spite of the disturbances in the feed flow (changes in C_{in} and/or T_{in}), while satisfying the process constraints.

The spatial domain was divided into $N = 300$ sections ➔ The linear model has **600** states.

Derivation of the reduced-order model of the reactor using POD

In POD, we start by observing that $\mathbf{x}(t) \in \mathbb{R}^{2N}$ can be expanded as a sum of orthonormal basis vectors :

$$\mathbf{x}(t) = \sum_{j=1}^{2N} a_j(t)\varphi_j$$

$$\varphi_j \in \mathbb{R}^{2N} \ , \ a_j(t) \in \mathbb{R}$$

$\{\varphi_j\}_{j=1}^{2N}$ → POD basis vectors

$\{a_j(t)\}_{j=1}^{2N}$ → POD coefficients (time-varying)

The main dynamics of the system can be represented using only the first n most relevant basis vectors.

$$\mathbf{x}_n(t) = \sum_{j=1}^{n} a_j(t)\varphi_j, \quad n \ll 2N$$

nth order approximation of $\mathbf{x}(t)$

By building a dynamic model for the first n POD coefficients we can derive a reduced order model for the system.

This is the essence of model reduction by POD

Derivation of the reduced-order model of the reactor using POD

The model was derived in 4 steps.

1. Generation of the snapshot matrix.

A snapshot matrix $\mathbf{X}_{\text{snap}} \in \mathbb{R}^{600\times1500}$ was constructed from the system response when independent step changes were applied to $\mathbf{u}(t)$ and $\mathbf{d}(t)$.

$$\mathbf{X}_{\text{snap}} = \left[\mathbf{x}(t=\Delta t), \mathbf{x}(t=2\Delta t), \ldots, \mathbf{x}(t=1500\Delta t)\right] \qquad \Delta t = 0.05\,\text{s}$$

2. Derivation of the POD basis vectors

They were derived by calculating the SVD of $\mathbf{X}_{\text{snap}}$,

Basis vectors

$$\mathbf{X}_{\text{snap}} = \mathbf{\Phi\Sigma\Psi}^T \qquad \mathbf{\Phi} \in \mathbb{R}^{600\times600} = \underbrace{\left[\varphi_1, \varphi_2, \ldots, \varphi_{600}\right]}_{\text{Basis vectors}}$$

3. Selection of the most relevant POD basis vectors

It was done by using the so-called energy criterion,

$$\overline{P}_n = \frac{\sum_{j=1}^{n} \sigma_j^2}{\sum_{j=1}^{2N} \sigma_j^2}, \qquad n = 1, \ldots, 2N$$

$\overline{P}_n =$ truncation degree of the selected basis vectors

Derivation of the reduced-order model of the reactor using POD

The first 20 POD basis vectors were selected. So, the 20th order approximation of $\mathbf{x}(t)$ is as follows:

$$\mathbf{x}_n(t) = \sum_{j=1}^{20} a_j(t)\varphi = \mathbf{\Phi}_n \mathbf{a}(t)$$

$$\mathbf{\Phi}_n = \left[\varphi_1, \varphi_2, \ldots, \varphi_{20}\right] \quad \mathbf{a}(t) = \left[a_1(t), a_2(t), \ldots, a_{20}(t)\right]^T$$

4. Construction of the model for the first $n = 20$ POD coefficients.

It was built by means of the Galerkin Projection. So, the linear model of the reactor was projected into the space spanned by the selected POD basis vectors.

The reduced order model (with only 20 states) is:

$$\dot{\mathbf{a}}(t) = \mathbf{A}_\mathrm{r}\mathbf{a}(t) + \mathbf{B}_\mathrm{r}\mathbf{u}(t) + \mathbf{F}_\mathrm{r}\mathbf{d}(t)$$
$$\mathbf{x}_n(t) = \mathbf{\Phi}_n \mathbf{a}(t)$$

Bilinear Transformation →

discrete-time version

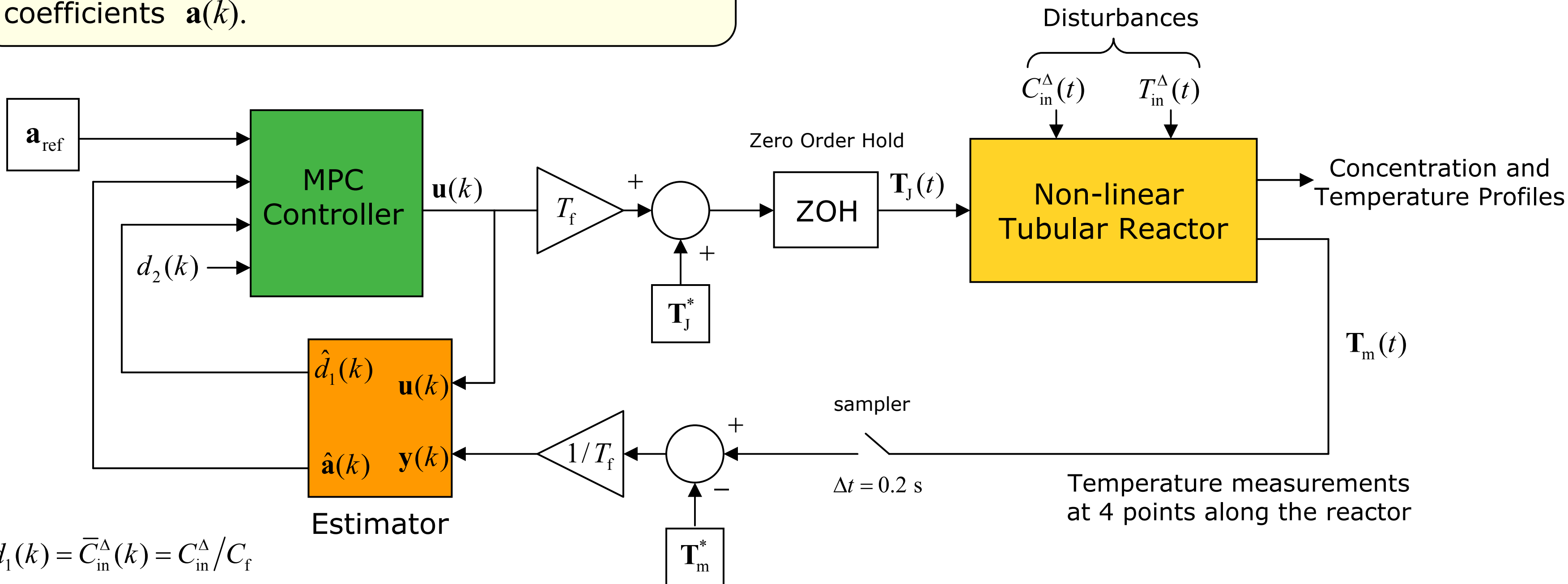
$$\mathbf{a}(k+1) = \tilde{\mathbf{A}}\mathbf{a}(k) + \tilde{\mathbf{B}}\mathbf{u}(k) + \tilde{\mathbf{F}}\mathbf{d}(k)$$
$$\mathbf{x}_n(k) = \mathbf{\Phi}_n \mathbf{a}(k)$$

where : $\mathbf{A}_\mathrm{r} = \mathbf{\Phi}_n^T \mathbf{A} \mathbf{\Phi}_n$, $\mathbf{B}_\mathrm{r} = \mathbf{\Phi}_n^T \mathbf{B}$ and $\mathbf{F}_\mathrm{r} = \mathbf{\Phi}_n^T \mathbf{F}$

$T_\mathrm{s} = 0.2\,\mathrm{s}$

MPC control scheme

The temperature and concentration profiles are controlled indirectly by controlling the POD coefficients $\mathbf{a}(k)$.



$$d_1(k) = \bar{C}_{\text{in}}^{\Delta}(k) = C_{\text{in}}^{\Delta} / C_{\text{f}}$$

$$d_2(k) = \bar{T}_{\text{in}}^{\Delta}(k) = T_{\text{in}}^{\Delta} / T_{\text{f}}$$

* → Values at the Operating Point

The references of the POD coefficients were calculated using this relation: $\mathbf{a}_{\text{ref}} = \mathbf{\Phi}_n^T \mathbf{x}_{\text{ref}}$

Since the control system has to keep the reactor around the operating profiles, the reference for $\mathbf{x}(k)$ is $\mathbf{0}$. ➡ $\mathbf{a}_{\text{ref}} = \mathbf{0}$

POD-based MPC controller for the reactor (MPC-QP)

The MPC formulation is as follows:

$$\min_{\mathbf{a}_{N_p},\Delta\mathbf{u}_{N_c},\mathbf{d}_{N_p},\xi} \sum_{i=1}^{N_p} \left\|\mathbf{a}_{\mathrm{ref}}(k+i)-\mathbf{a}(k+i)\right\|_{\mathbf{Q}}^2 + \sum_{i=0}^{N_c-1} \left\|\Delta\mathbf{u}(k+i)\right\|_{\mathbf{R}}^2 + P_{\mathrm{Q}}\xi^2 + P_{\mathrm{L}}\xi$$

Subject to:

$$\begin{aligned}
&\mathbf{a}(k+i+1) = \tilde{\mathbf{A}}\mathbf{a}(k+i)+\tilde{\mathbf{B}}\mathbf{u}(k+i)+\tilde{\mathbf{F}}\mathbf{d}(k+i), && i=0,\ldots,N_{\mathrm{p}}-1\\
&\mathbf{d}(k+i+1) = \mathbf{d}(k+1), && i=0,\ldots,N_{\mathrm{p}}-1\\
&\mathbf{u}(k+i) = \mathbf{u}(k+i-1)+\Delta\mathbf{u}(k+i), && i=0,\ldots,N_{\mathrm{c}}-1\\
&\mathbf{u}(k+i) = \mathbf{u}(k+i-1), && i=N_{\mathrm{c}},\ldots,N_{\mathrm{p}}-1\\
&\mathbf{u}_{\min} \le \mathbf{u}(k+i) \le \mathbf{u}_{\max}, && i=0,\ldots,N_{\mathrm{c}}-1\\
&\overline{\mathbf{T}}^{\Delta}(k+i) = \mathbf{\Phi}_{\mathrm{T}}\mathbf{a}(k+i) \le \overline{\mathbf{T}}^{\Delta_{\max}} + \mathbf{1}\cdot\eta(i)\xi, && i=1,\ldots,N_{\mathrm{p}}\\
&\xi \ge 0,
\end{aligned}$$

where : $\mathbf{\Phi}_{20} = \left[\mathbf{\Phi}_{\mathrm{C}}^T, \mathbf{\Phi}_{\mathrm{T}}^T\right]^T$, $\eta(i) = 1/r_{\mathrm{c}}^{i-1}$, $r_{\mathrm{c}} > 1$, $\overline{\mathbf{T}}^{\Delta_{\max}} = \left(400\ \mathrm{K}\cdot\mathbf{1}-\mathbf{T}^*\right)/T_{\mathrm{f}}$

Mechanism for handling infeasibilities: a slack variables approach (L_∞-norm and time-dependent weights)

Parameters: $\mathbf{Q}=\mathbf{I}_{20}$, $\mathbf{R}=110\cdot\mathrm{I}_3$, $r_{\mathrm{c}}=1.2$, $P_{\mathrm{L}}=P_{\mathrm{Q}}=10^4$, $N_{\mathrm{c}}=10$, $N_{\mathrm{p}}=80$

Simulation Results: Test 1

T_{in} and C_{in} are increased by $10\ \mathrm{K}$ and 10^{-3} mol/l.

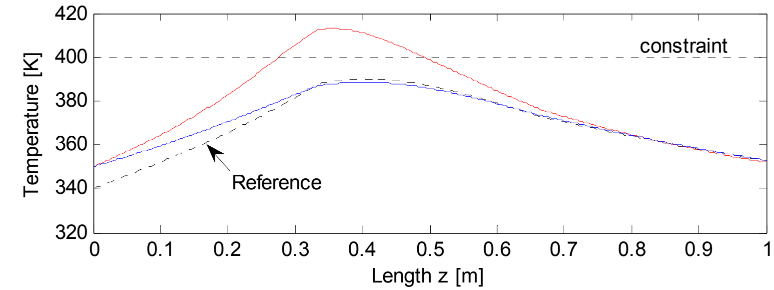
Quantities of interest

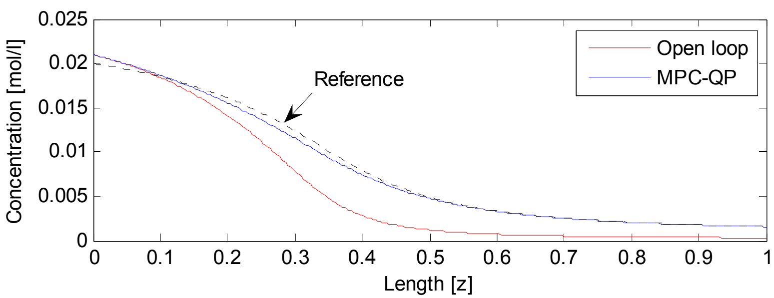
Control	T_{max} [K]	ΔC_{out} [%]
Open loop	413.03	-79.39
MPC-QP	393.94	-0.67

T_{max} ➔ Maximum temperature reached inside the reactor during the test.

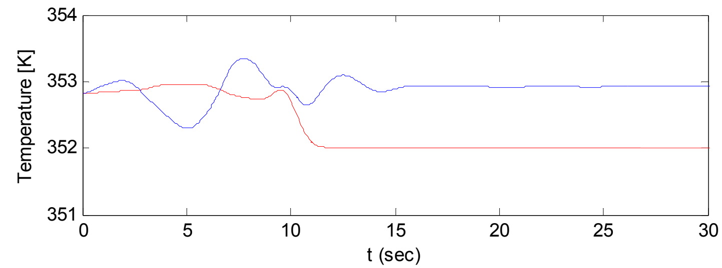
ΔC_{out} ➔ Percentage of change of the concentration at the reactor output with respect to its nominal value (steady state).

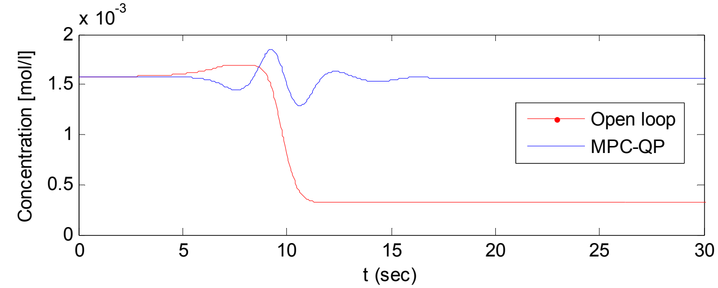
Steady-state profiles of the reactor



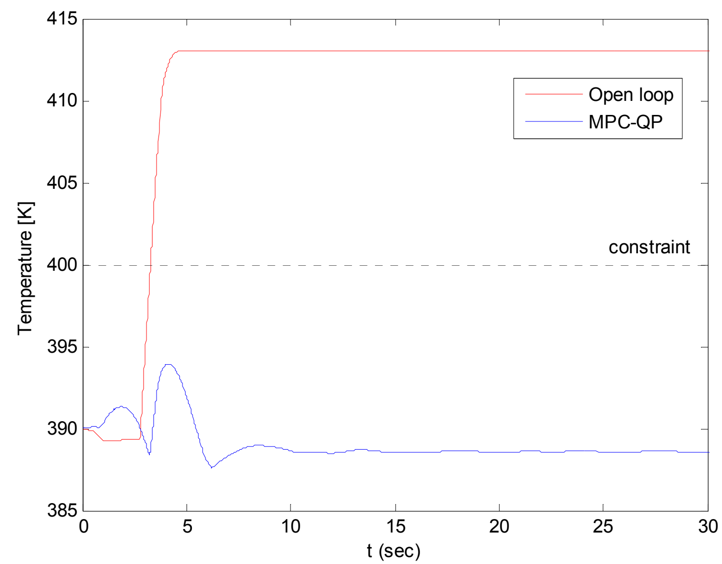


Temperature and concentration at the reactor outlet





Maximum Peak of the temperature profile



Simulation Results: Test 2

T_{in} and C_{in} are decreased by $10\ K$ and 10^{-3} mol/l.

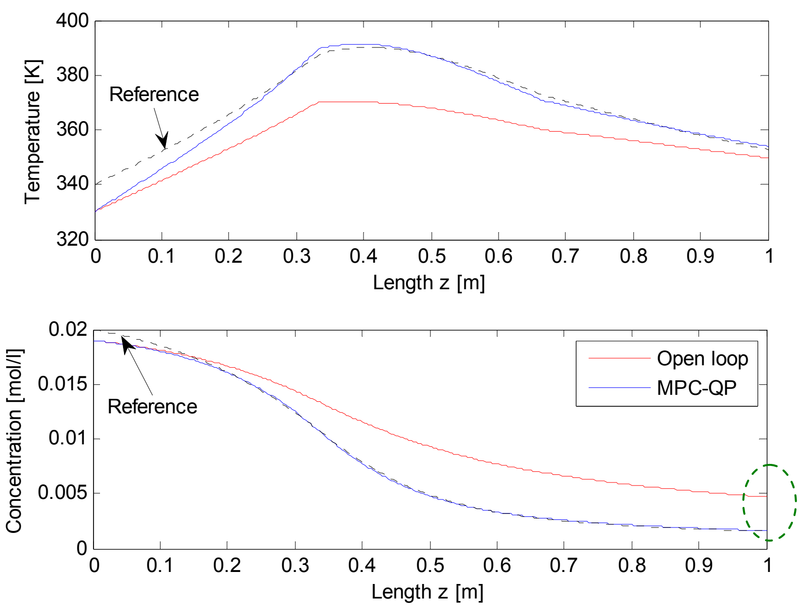
Quantities of interest

Control	T_{max} [K]	ΔC_{out} [%]
Open loop	390	198.25
MPC-QP	397.58	3.20

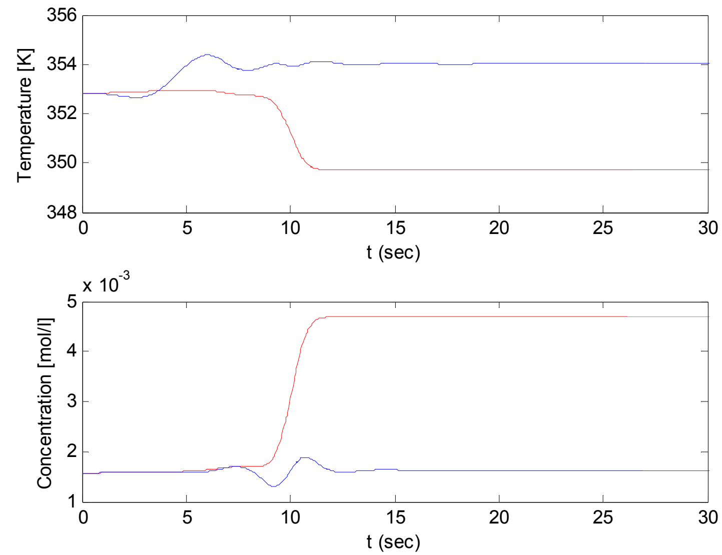
T_{max} ➔ Maximum temperature reached inside the reactor during the test.

ΔC_{out} ➔ Percentage of change of the concentration at the reactor output with respect to its nominal value (steady state).

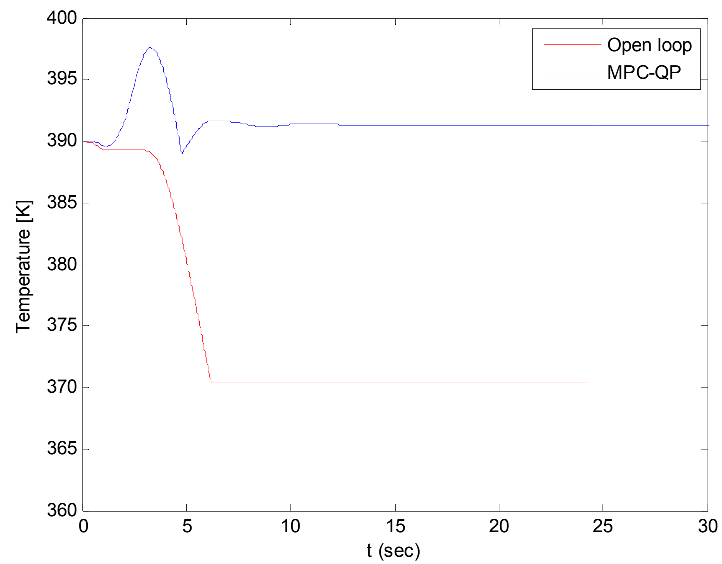
Steady-state profiles of the reactor



Temperature and concentration at the reactor outlet



Maximum Peak of the temperature profile



Number of state/output constraints of MPC-QP

Problem:

Temperature constraints: $\bar{\mathbf{T}}^{\Delta}(k+i) = \mathbf{\Phi}_{\mathrm{T}}\mathbf{a}(k+i) \leq \bar{\mathbf{T}}^{\Delta_{\max}} + \mathbf{1}\cdot\eta(i)\xi, \qquad i = 1,\ldots,N_{\mathrm{p}}$

Number of temperature constraints: $N \times N_{\mathrm{p}} = 300 \times 80 = 24000$

The MPC-QP controller has to deal with a large number of temperature constraints which demand a considerable amount of computational resources.

Resolution Methods:

- Positive Polynomial Approach.
- Greedy Selection Algorithm.

Positive Polynomial Approach

Approximation of the temperature constraints

If we use univariate polynomials ($P_{\max}(z)$ and $P_j(z)$) of degree d for approximating $\overline{\mathbf{T}}^{\Delta_{\max}}$ and the part of the basis vectors associated to the temperature profile $\mathbf{\Phi}_{\mathrm{T}} = [\tilde{\varphi}_1, \tilde{\varphi}_2, \ldots, \tilde{\varphi}_{20}]$, we can approximate the temperature constraints ($\mathbf{\Phi}_{\mathrm{T}}\mathbf{a}(k) \leq \overline{\mathbf{T}}^{\Delta_{\max}}$) as follows:

$$\sum_{j=1}^{n=20} a_j(k) P_j(z) \leq P_{\max}(z), \ \forall z \in [0, \ 1] \quad \Rightarrow \quad \begin{cases} P^{(k)}(z) = P_{\max}(z) - \sum_{j=1}^{n=20} a_j(k) P_j(z) \\ P^{(k)}(z) \geq 0, \quad \forall z \in [0, \ 1] \end{cases} \qquad (1)$$

Semidefinite representability of positive polynomials

Univariate real polynomials are nonnegative everywhere iff they can be written as a sum of squared polynomials. This property is denoted by the acronym SOS (Sum Of Squares).

Proposition 1: A univariate polynomial $P(z)$ of degree $2d$ is SOS iff there exists a $(d+1)\times(d+1)$ positive semidefinite matrix $\mathbf{W}$ such that

$$P(z) = \mathbf{f}(z)^T \mathbf{W} \mathbf{f}(z),$$

where $\mathbf{f}(z) = \left[1, z, z^2, \ldots, z^d\right]^T$.

Positive Polynomial Approach

It is possible to relate the positivity of a real univariate polynomial on a compact interval to the positivity of other polynomial on the whole real line by the following transformation.

Proposition 2: A real univariate polynomial p of degree d is nonnegative on the compact interval $[a, b]$ iff

$$(1+z^2)^d\, p\left(a + \frac{(b-a)z^2}{1+z^2}\right) \geq 0, \quad \forall z \in \mathbb{R}.$$

For every $1 \leq k \leq N_{\mathrm{p}}$, the condition (1) can be converted into:

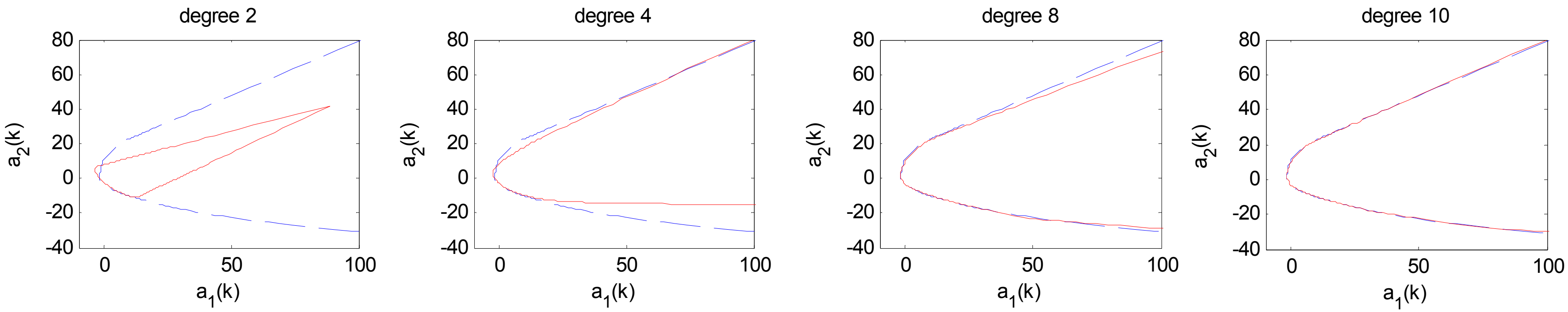
$$P^{(k)}(z) \geq 0, \quad \forall z \in [0,\ 1] \quad \Longrightarrow \quad \tilde{P}^{(k)}(z) = (1+z^2)^d P^{(k)}\left(\frac{z^2}{1+z^2}\right) \geq 0, \quad \forall z \in \mathbb{R},$$

And, denoting by $\mathbf{S}_+^{d+1}$ the set of $(d+1)\times(d+1)$ positive semidefinite matrices, into

Semidefinite feasibility problem

$$\begin{aligned} &\text{find} \quad \mathbf{W}^{(k)} \in S_+^{d+1} \\ &\text{such that} \quad \tilde{P}^{(k)}(z) = \mathbf{f}(z)^T \mathbf{W}^{(k)} \mathbf{f}(z). \end{aligned} \qquad (2)$$

Positive Polynomial Approach



Feasible regions delimited by the temperature constraints of a 2nd order POD model. Blue line – Full set of constraints. Red line – Polynomial approximation given by (2).

In the formulation of the MPC based on the polynomial approximations (MPC-SDP), the inequality constraints are replaced by

$$\mathbf{W}^{(k+i)} \succeq 0 \qquad \tilde{P}^{(k+i)}(z) = \mathbf{f}(z)^T \mathbf{W}^{(k+i)} \mathbf{f}(z)$$

with

$$\begin{cases} \tilde{P}^{(k+i)}(z) = (1+z^2)^d \, P^{(k+i)}\left(z^2/(1+z^2)\right) \\ P^{(k+i)}(z) = P_{\max}(z) + \eta(i)\xi - \sum_{j=1}^{n=20} a_j(k+i) P_j(z) \end{cases}$$

MPC-QP:

24 000 inequality constraints

MPC-SDP: $d = 12$

$(2d+1) \times N_p = 2000$ Equality constraints

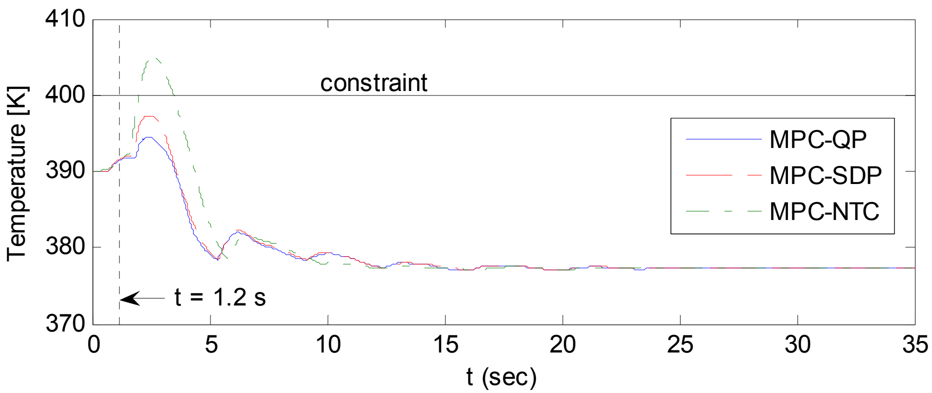
$N_p = 80$ LMIs of dimension 13 × 13

Positive Polynomial Approach

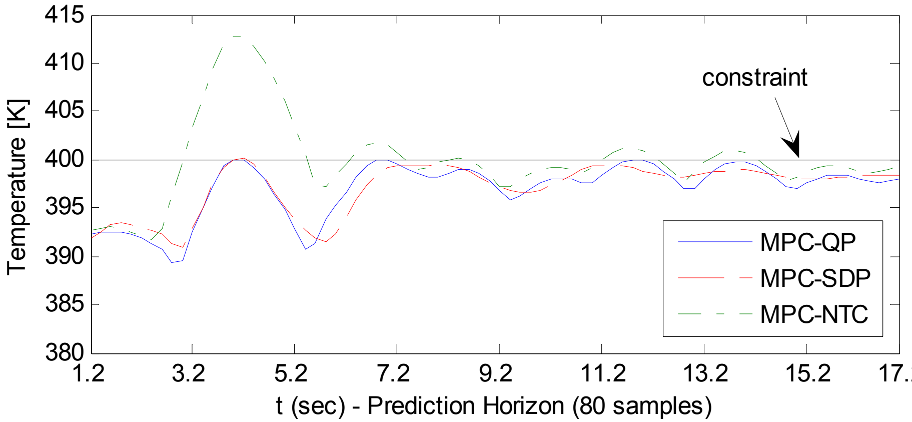
Some simulation results

Test 3: T_{in} and C_{in} are increased by 24 K and $3{\cdot}10^{-3}$ mol/l with respect to their nominal values.

Maximal peak of the temperature profile



Predictions of the maximal peak of the temperature profile at $t = 1.2$ s

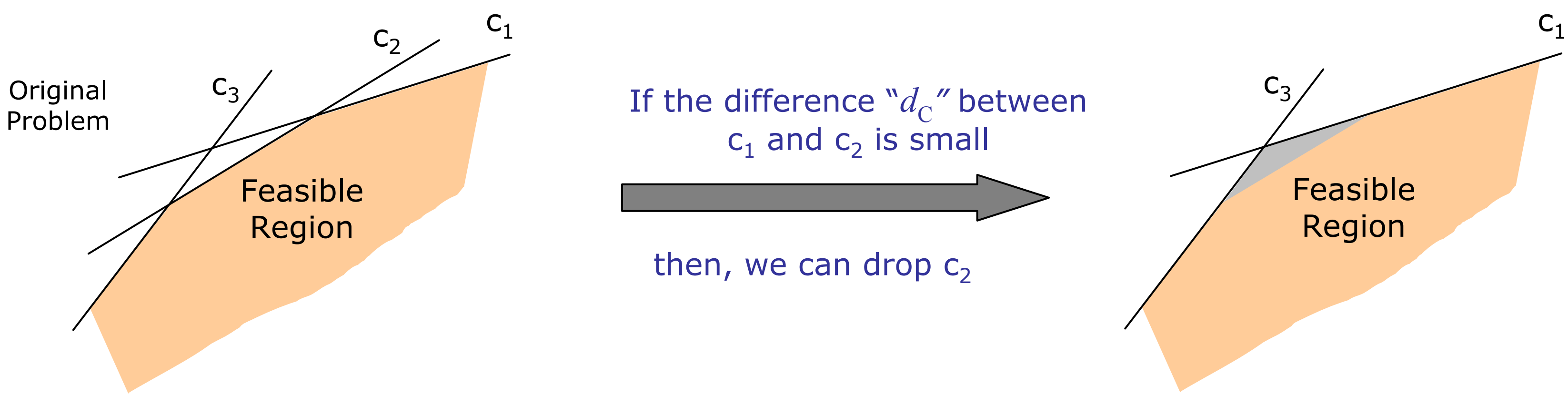


Control	No. opt. variables	Number of constraints				Memory (MB)	Average time for solving the optimization
		Inequalities	Equalities	SOC	LMI		
MPC-QP (solver:Sedumi)	32	24061	-	1	-	6.2	13.01 s
MPC-QP (solver:Quadprog)	31	24061	-	-	-	6.02	0.31 s
MPC-SDP (solver: Sedumi)	7378	61	2065	2	80	0.67 (it requires 9 times less memory)	5.77 s

Positive polynomial approach → The temperature constraint is imposed at every point of the reactor (infinite number of ineq. constraints)

Greedy Selection Algorithm

Observation: "The coefficients of consecutive constraints are quite similar" ➔ By exploiting this, we can find a subset of constraints that approximates the feasible region of the complete set.



$c_1: \ \mathbf{\Phi}_{\mathrm{T}}(1,1)a_1(k) + \mathbf{\Phi}_{\mathrm{T}}(1,2)a_2(k) \le \overline{\mathbf{T}}^{\Delta_{\max}}(1)$

$c_2: \ \mathbf{\Phi}_{\mathrm{T}}(2,1)a_1(k) + \mathbf{\Phi}_{\mathrm{T}}(2,2)a_2(k) \le \overline{\mathbf{T}}^{\Delta_{\max}}(2)$

$c_3: \ \mathbf{\Phi}_{\mathrm{T}}(3,1)a_1(k) + \mathbf{\Phi}_{\mathrm{T}}(3,2)a_2(k) \le \overline{\mathbf{T}}^{\Delta_{\max}}(3)$

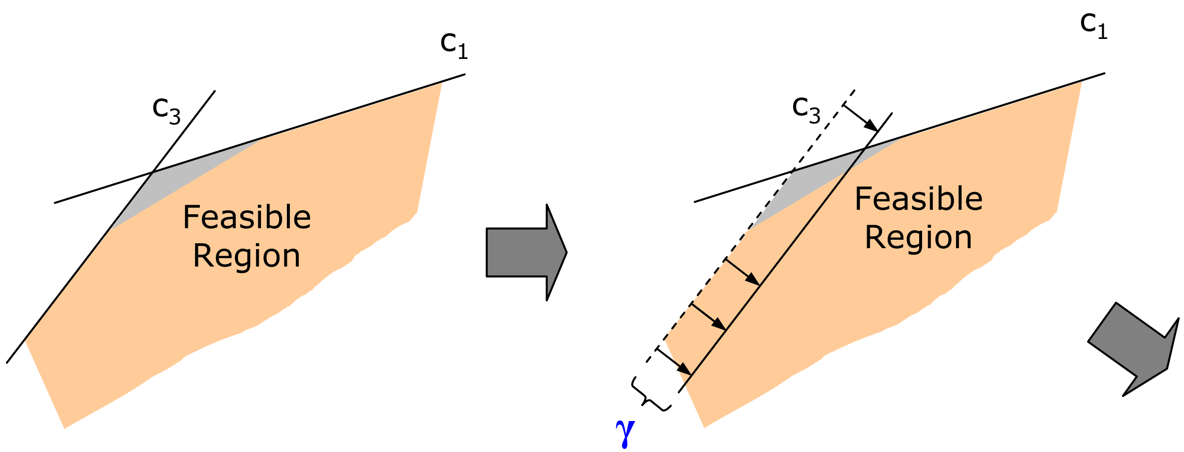
"d_{C}" is defined as the "mean absolute error" between the coefficients of the constraints that are being compared (p and i):

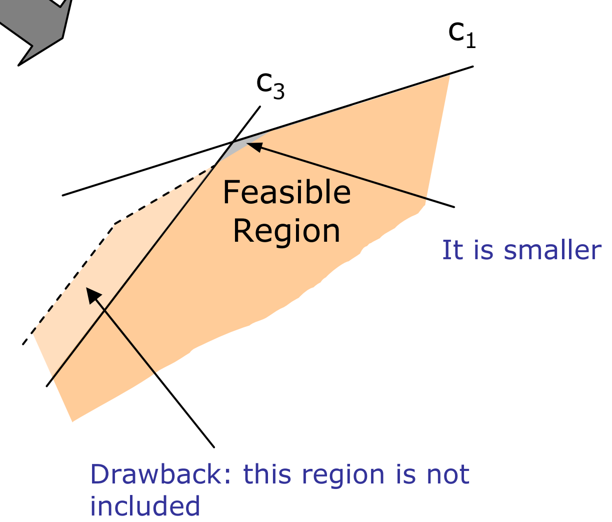
$$d_{\mathrm{C}}(p,i) = \frac{1}{n+1}\left(\sum_{j=1}^{n} \left| \mathbf{\Phi}_{\mathrm{T}}(p,j) - \mathbf{\Phi}_{\mathrm{T}}(i,j) \right| + \left| \overline{\mathbf{T}}^{\Delta_{\max}}(p) - \overline{\mathbf{T}}^{\Delta_{\max}}(i) \right| \right)$$

$n = 20 =$ number of POD coefficients

Greedy Selection Algorithm

In order to reduce the size of the area that does not belong to the original feasible region, we introduce the "shrinking" parameter γ to tighten non consecutive constraints ➔ conservative measure.





$$c_3 : \mathbf{\Phi}_T(3,1)a_1(k) + \mathbf{\Phi}_T(3,2)a_2(k) \leq \overline{\mathbf{T}}^{\Delta_{max}}(3)$$

$$c_3 : \mathbf{\Phi}_T(3,1)a_1(k) + \mathbf{\Phi}_T(3,2)a_2(k) \leq \overline{\mathbf{T}}^{\Delta_{max}}(3) - \gamma$$

Greedy Selection Algorithm

1) Set p=1, and select the first constraint:

$$\mathbf{T}_{\mathrm{R}} = \overline{\mathbf{T}}^{\Delta_{\max}}(1), \quad \mathbf{\Phi}_{\mathrm{R}} = \mathbf{\Phi}_{\mathrm{T}}(1,:)$$

2) For all $i = 2,\ldots,N-1$, perform

a) Calculate the difference between the pth and ith constraints: $d_{\mathrm{C}}(p,i)$

b) If $d_{\mathrm{C}}(p,i) \geq Sel$ then select the ith constraint:

- $\mathbf{\Phi}_{\mathrm{R}} = \left[\mathbf{\Phi}_{\mathrm{R}};\mathbf{\Phi}_{\mathrm{T}}(i,:)\right]$
- If $(i\text{-}p) > 1$ then $\mathbf{T}_{\mathrm{R}} = \left[\mathbf{T}_{\mathrm{R}};\overline{\mathbf{T}}^{\Delta_{\max}}(i) - \gamma\right]$

  else $\mathbf{T}_{\mathrm{R}} = \left[\mathbf{T}_{\mathrm{R}};\overline{\mathbf{T}}^{\Delta_{\max}}(i)\right]$
- Set $p=i$

3) Select the last constraint:

$$\mathbf{T}_{\mathrm{R}} = \left[\mathbf{T}_{\mathrm{R}};\overline{\mathbf{T}}^{\Delta_{\max}}(N)\right], \quad \mathbf{\Phi}_{\mathrm{R}} = \left[\mathbf{\Phi}_{\mathrm{R}};\mathbf{\Phi}_{\mathrm{T}}(N,:)\right]$$

Input

$$\mathbf{\Phi}_{\mathrm{T}}\mathbf{a}(k) \leq \overline{\mathbf{T}}^{\Delta_{\max}}$$

$$\mathbf{\Phi}_{\mathrm{T}} \in \mathbb{R}^{N\times n}, \quad \overline{\mathbf{T}}^{\Delta_{\max}} \in \mathbb{R}^{N}$$

Output

$$\mathbf{\Phi}_{\mathrm{R}}\mathbf{a}(k) \leq \mathbf{T}_{\mathrm{R}}$$

$$\mathbf{\Phi}_{\mathrm{R}} \in \mathbb{R}^{S_c\times 20}, \quad \mathbf{T}_{\mathrm{R}} \in \mathbb{R}^{S_c}$$

$$S_c \ll N$$

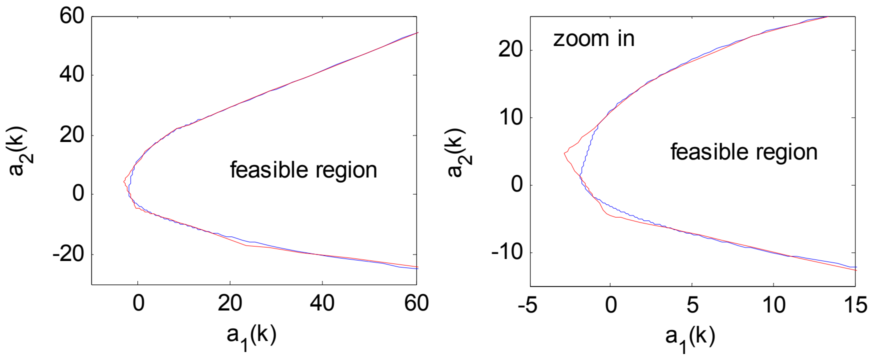
Parameters of the algorithm

Sel = Threshold for selecting a constraint

γ = Shrinking parameter

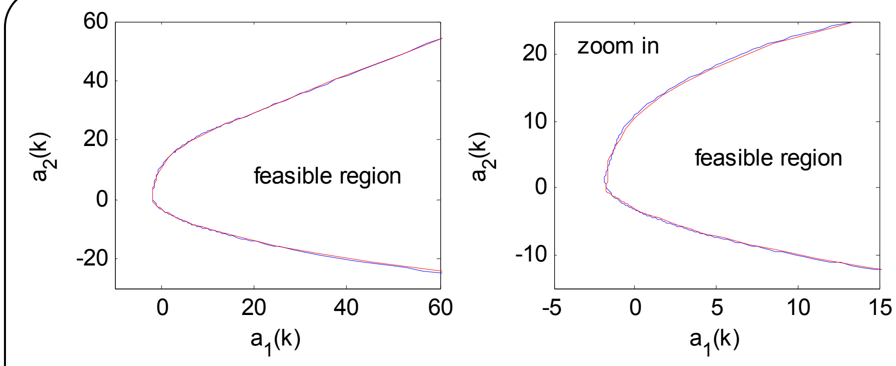
Greedy Selection Algorithm

Blue Line – Full set of constraints (300). Red Line – Reduced set of constraints.



Number of selected constraints: **7**

Parameters : $Sel = 0.08, \quad \gamma = 0.01 \quad (0.8\ \mathrm{K} / T_\mathrm{f})$



Number of selected constraints: **21**

Parameters : $Sel = 0.022, \quad \gamma = 0.01 \quad (0.8\ \mathrm{K} / T_\mathrm{f})$

MPC-QP-RS ➔ MPC with a reduced set of constraints

MPC-QP

$$\mathbf{\Phi}_\mathrm{T}\mathbf{a}(k+i) \le \overline{\mathbf{T}}^{\Delta_{\max}} + \mathbf{1}\cdot\eta(i)\xi$$

$$\mathbf{\Phi}_\mathrm{T} \in \mathbb{R}^{300\times 20}, \quad \overline{\mathbf{T}}^{\Delta_{\max}} \in \mathbb{R}^{300}$$

24 000 Temperature constraints

Algorithm

$Sel = 0.03$

$\gamma = 0.00625\ (0.5\ \mathrm{K})$

MPC-QP-RS

$$\mathbf{\Phi}_\mathrm{R}\mathbf{a}(k+i) \le \mathbf{T}_\mathrm{R} + \mathbf{1}\cdot\eta(i)\xi$$

$$\mathbf{\Phi}_\mathrm{R} \in \mathbb{R}^{S_\mathrm{c}\times 20}, \quad \mathbf{T}_\mathrm{R} \in \mathbb{R}^{S_\mathrm{c}} \qquad S_\mathrm{c} = 20$$

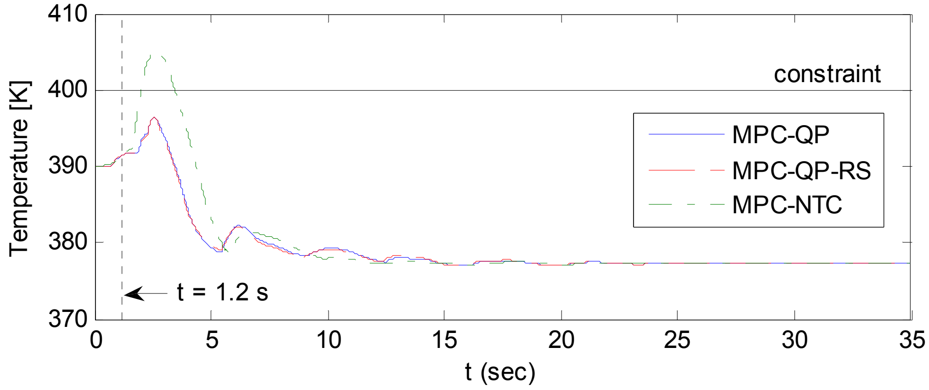
1600 Temperature constraints

Greedy Selection Algorithm

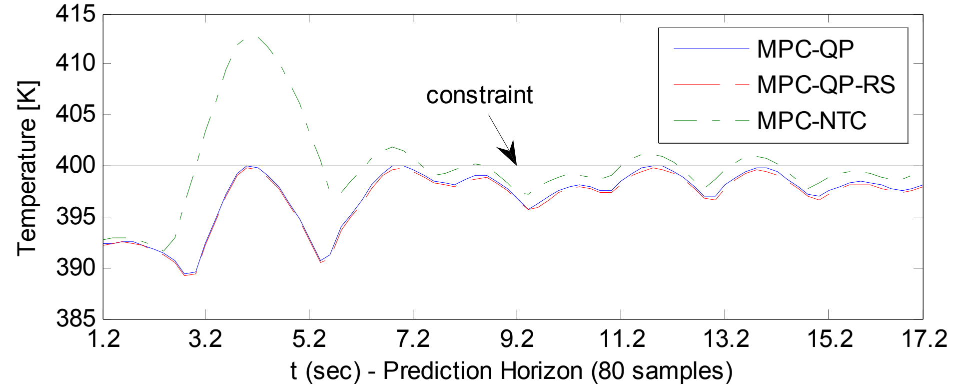
Some simulation results

Test 3: T_{in} and C_{in} are increased by $24\ K$ and $3{\cdot}10^{-3}$ mol/l with respect to their nominal values.

Maximal peak of the temperature profile



Predictions of the maximal peak of the temperature profile at $t = 1.2$ s



Control	No. opt. Variables	No. inequality Constraints	Memory (MB)	Average time for solving the optimization
MPC-QP	31	24061	6.02	0.31 s
MPC-QP-RS	31	1661	0.42 (it requires 14.33 times less memory)	0.023 s (it is solved 13.48 times faster)

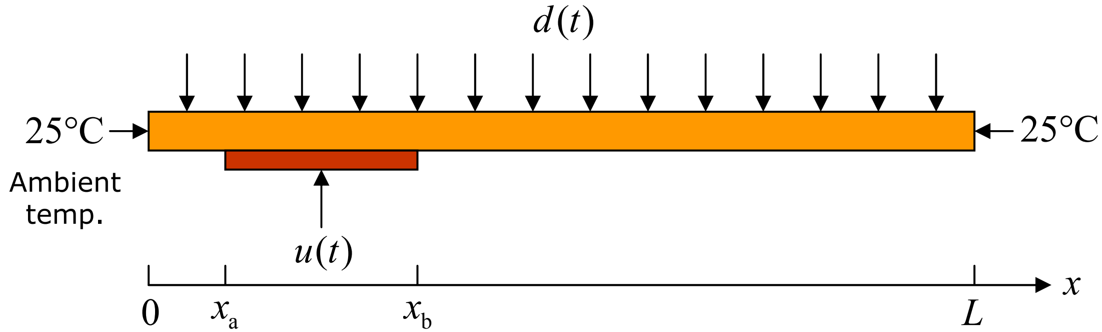
ESAT

PART II

Acceleration of the evaluation of nonlinear POD models

Nonlinear heat transfer in a one-dimensional bar

System under consideration :



$$\left.\begin{matrix} d(t) \\ u(t) \end{matrix}\right\} \text{Heat Fluxes } [\mathrm{W}\cdot\mathrm{m}^{-3}]$$

If only temperature variations in the x-direction are considered, the dynamics of temperature $T(x,t)$ of the bar can be modeled by

$$\rho C_{\mathrm{p}} \frac{\partial T(x,t)}{\partial t} = \frac{\partial}{\partial x}\left(\kappa\big(T(x,t)\big) \frac{\partial T(x,t)}{\partial x} \right) + V(x,t)$$

where

$$V(x,t) = \begin{cases} d(t) + u(t), & x_a \le x \le x_b \\ d(t), & \text{elsewhere} \end{cases}$$

Heat Conductivity:

$$\kappa(T) = \kappa_0 + \kappa_1 T + \kappa_2 T^2 + \kappa_3 T^3$$

Initial Condition:

$$T(x,0) = 25°\mathrm{C}$$

Boundary Conditions (Direchlet):

$$T(x=0,t) = T(x=L,t) = 25°\mathrm{C}$$

Nonlinear heat transfer in a one-dimensional bar

Discretized model:

$$\dot{\mathbf{T}}(t) = \mathbf{F}\left(\mathbf{T}(t)\right) + \mathbf{B}_1 d(t) + \mathbf{B}_2 u(t)$$

$$\mathbf{T}(t) \in \mathbb{R}^{N-1} = \left[T_1(t), T_2(t), \ldots, T_{N-1}(t)\right]$$

$$\mathbf{F}\left(\mathbf{T}(t)\right) : \mathbb{R}^{N-1} \rightarrow \mathbb{R}^{N-1},\ N = 500 \text{ sections}$$

Nonlinear POD model:

$$\dot{\mathbf{a}}(t) = \underbrace{\mathbf{\Phi}_n^T \mathbf{F}\left(\mathbf{\Phi}_n \mathbf{a}(t) + \mathbf{T}^*\right)}_{\mathbf{f}\left(\mathbf{a}(t)\right) : \mathbb{R}^6 \rightarrow \mathbb{R}^6} + \tilde{\mathbf{B}}_1 d(t) + \tilde{\mathbf{B}}_2 u(t)$$

$$\mathbf{a}(t) \in \mathbb{R}^6$$

$$\mathbf{T}_n(t) = \mathbf{\Phi}_n \mathbf{a}(t) + \mathbf{T}^*$$

It is required the high-dimensional vector function $\mathbf{F}$!!!

Problem: The evaluation of the derivatives is COMPUTATIONALLY INTENSIVE !!!!

Resolution Methods:

- Neural Network Approach.
- Polynomial-POD model approach.

Neural Network Approach

Idea: To approximate the function $\mathbf{y}(t) = \mathbf{f}(\mathbf{a}(t))$ by an MLP.

Justification:
- An MLP can learn any nonlinear input-output mapping
- The evaluation of a trained MLP can be done very fast

Structure of the MLP used

- Inputs : 6
- Hidden neurons (Hyperbolic tangent) : 10
- Output neurons (Linear) : 6

Training Details

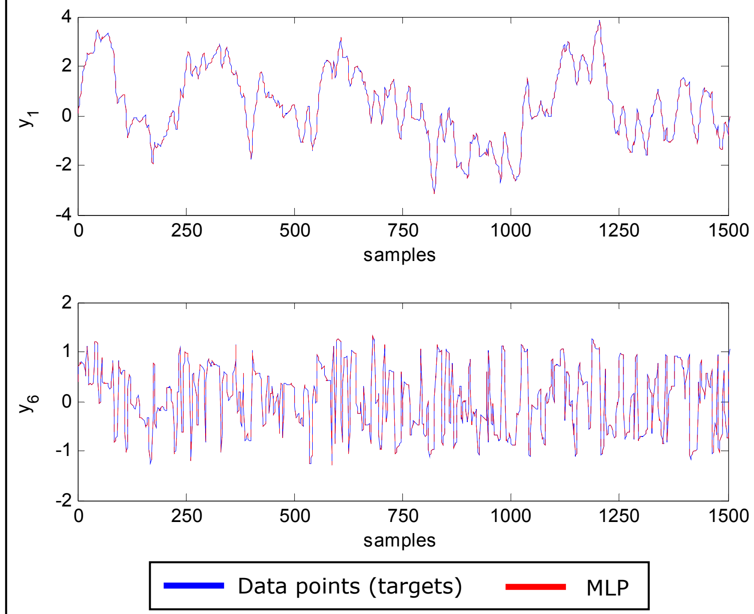
- Algorithm: Levenberg-Marquardt.
- For avoiding Overfitting: Early Stopping criterion.
- Training epochs: 6000

Neural-POD model

$$\dot{\mathbf{a}}(t) = \hat{\mathbf{f}}\left(\mathbf{a}(t)\right) + \tilde{\mathbf{B}}_1 d(t) + \tilde{\mathbf{B}}_2 u(t)$$

$$\hat{\mathbf{f}}(\mathbf{a}(t)) = \mathbf{v}^{-1}\left(\mathbf{W}^{\mathrm{o}} \cdot \mathbf{g}\left(\mathbf{W}^{\mathrm{h}} \cdot \mathbf{h}\left(\mathbf{a}(t)\right) + \mathbf{b}^{\mathrm{h}}\right) + \mathbf{b}^{\mathrm{o}}\right)$$

MLP test performance for $y_1(t)$ and $y_6(t)$



MSE Training set: 8.0662×10^{-7}

MSE Test set: 8.0975×10^{-7}

Polynomial-POD model Approach

Idea: To exploit the polynomial nature of $\mathbf{y}(t) = \mathbf{f}(\mathbf{a}(t))$ in order to find an alternative representation that can be evaluated much faster.

P-POD model

By expanding and simplifying

$$\mathbf{f}\left(\mathbf{a}(t)\right) = \mathbf{\Phi}_n^T \mathbf{F}\left(\mathbf{\Phi}_n \mathbf{a}(t) + \mathbf{T}^*\right),$$

we get a compact representation of $\mathbf{f}(\mathbf{a}(t))$.

"Multivariate polynomials in terms of $a_j(t), \forall j = 1,\ldots,n$"

$$\dot{\mathbf{a}}(t) = \mathbf{f}\left(\mathbf{a}(t)\right) + \tilde{\mathbf{B}}_1 d(t) + \tilde{\mathbf{B}}_2 u(t)$$

$$\mathbf{f}(\mathbf{a}(t)) = \left[f_1\left(\mathbf{a}(t)\right),\ldots, f_m\left(\mathbf{a}(t)\right),\ldots, f_n\left(\mathbf{a}(t)\right)\right]^T$$

$$f_m\left(\mathbf{a}(t)\right) = w_{m,0} + w_{m,1} a_1(t) + \cdots + w_{m,n} a_n(t) + w_{m,(n+1)} a_1^2(t) + w_{m,(n+2)} a_1(t) a_2(t) + \cdots + w_{m,(r-1)} a_n^{d_p}(t)$$

The coefficients of the polynomials are calculated by solving the following least squares problem (fitting polynomials to data):

$$\min_{\mathbf{w}} J = \left(\mathbf{y}_{N_d} - \hat{\mathbf{y}}_{N_d}\right)^T \left(\mathbf{y}_{N_d} - \hat{\mathbf{y}}_{N_d}\right)$$

data — predictions

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\left(\mathbf{\Omega}^T\mathbf{\Omega}\right)\mathbf{w} - \left(\mathbf{y}_{N_d}^T \mathbf{\Omega}\right)\mathbf{w}$$

$\mathbf{w}$ → Coefficients of the Polynomials

$\mathbf{w} \in \mathbb{R}^{210*6=1260}$

Polynomial-POD model Approach

Limitation: the number of monomials per component function $r=\sum_{j=0}^{d_p}\frac{(n+j-1)!}{j!(n-1)!}$ increases exponentially with the number of POD coefficients (n).

Solution: To select a reduced set of monomials using a sequential feature selection method.

Sequential feature selection method

- Objective function to minimize (SSE): $J^{\mathcal{S}}=\left(\mathbf{y}_{N_d}-\hat{\mathbf{y}}^{\mathcal{S}}_{N_d}\right)^T\left(\mathbf{y}_{N_d}-\hat{\mathbf{y}}^{\mathcal{S}}_{N_d}\right)$

  data — Predictions of the candidate subset

- Sequential search algorithm: Sequential Forward Selection (SFS)

Evaluation of each candidate subset : a 10-fold cross-validation scheme is used.

Calculation of the coefficients : via Least squares

P-POD-RS model:

$$\dot{\mathbf{a}}(t)=\mathbf{f}^{\mathcal{S}^*}\left(\mathbf{a}(t)\right)+\tilde{\mathbf{B}}_1 d(t)+\tilde{\mathbf{B}}_2 u(t)$$

Each component function contains 25 monomials

Polynomial-POD model Approach

Polynomial-POD models with stability guarantee

The local stability of the P-POD model around the origin is analyzed by its autonomous counterpart :

$$\dot{\mathbf{a}}(t) = \breve{\mathbf{f}}\left(\mathbf{a}(t)\right) = \mathbf{f}\left(\mathbf{a}(t)\right)\Big|_{w_{1,0}=w_{2,0}=\cdots w_{n,0}=0}$$

$$u(t) = d(t) = 0$$

$$w_{1,0} = w_{2,0} = \cdots = w_{n,0} = 0$$

Lyapunov's indirect method ➔ the stability of the previous system is inferred from the stability of the linearized system

$$\delta\dot{\mathbf{a}}(t) = \mathbf{A}\delta\mathbf{a}(t)$$

➔ The origin is asymptotically stable if $\mathbf{A}$ is Hurwitz !!!

Non-convex, Non-smooth optimization problem:

$$\min_{\mathbf{w}} \; \frac{1}{2}\mathbf{w}^T\left(\mathbf{\Omega}^T\mathbf{\Omega}\right)\mathbf{w} - \left(\mathbf{y}_{N_d}^T\mathbf{\Omega}\right)\mathbf{w}$$

Subject to

$$\mathrm{Re}\left[\lambda_i(\mathbf{A}(\mathbf{w}))\right] < 0, \quad \text{for } i = 1,\ldots,n.$$

Jacobian

$$\mathbf{A} \in \mathbb{R}^{n\times n} = \frac{\partial\breve{\mathbf{f}}}{\partial\mathbf{a}}\Big|_{\mathbf{a}=0} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{bmatrix}$$

"Local stability around the origin depends only on the linear terms"

Polynomial-POD model Approach

Semidefinite Problem Formulation

Lemma: Let $\mathbf{A}$ be a square matrix. If the Hermitian part of $\mathbf{A}$, i.e. $0.5\ (\mathbf{A}+\mathbf{A}^{\mathrm{H}})$, is negative definite, then $\mathbf{A}$ is Hurwitz.

$$\min_{\mathbf{w}} \ \frac{1}{2}\mathbf{w}^T\left(\mathbf{\Omega}^T\mathbf{\Omega}\right)\mathbf{w}-\left(\mathbf{y}_{N_d}^T\mathbf{\Omega}\right)\mathbf{w}$$

subject to

$$-\frac{1}{2}\left(\mathbf{A}(\mathbf{w})+\mathbf{A}(\mathbf{w})^T\right)-\mu\mathbf{I}\succeq 0$$

$$\mu \geq 0$$

- Relaxation of the eigenvalue constraint
- Sufficient condition for the stability of the model ➔ It might be very conservative
- SDP problem (Convex) ➔ solved using *Sedumi*

Nonlinear Semidefinite Problem formulation

Theorem: Given $\delta\dot{\mathbf{a}}(t)=\mathbf{A}\delta\mathbf{a}(t)$, the origin is asymptotically stable iff, for any $\mathbf{Q}\succ 0$, there exists a $\mathbf{P}\succ 0$ such that

$$\mathbf{A}^T\mathbf{P}+\mathbf{P}\mathbf{A}+\mathbf{Q}=\mathbf{0}.$$

$$\min_{\mathbf{w},\mathbf{P}} \ \frac{1}{2}\mathbf{w}^T\left(\mathbf{\Omega}^T\mathbf{\Omega}\right)\mathbf{w}-\left(\mathbf{y}_{N_d}^T\mathbf{\Omega}\right)\mathbf{w}$$

subject to

$$\mathbf{A}(\mathbf{w})^T\mathbf{P}+\mathbf{P}\mathbf{A}(\mathbf{w})+\mathbf{Q}=\mathbf{0}$$

$$\mathbf{P}-\tilde{\mu}\mathbf{I}\succeq\mathbf{0}$$

$$\tilde{\mu}\geq 0$$

- Sufficient and necessary condition for the stability of the model.
- NSDP problem (Non-convex). It involves a BMI,

$$\mathbf{A}(\mathbf{w})^T\mathbf{P}+\mathbf{P}\mathbf{A}(\mathbf{w})+\tilde{\mu}\mathbf{I}\preceq\mathbf{0}$$

BMI solver ➔ *PEMBMI*

Some Simulation and Validation Results

Test 1: Steps of magnitude $1200 \cdot 10^3$ $W \cdot m^{-3}$ and $500 \cdot 10^3$ $W \cdot m^{-3}$ are applied to $u(t)$ and $d(t)$.

Performance of the POD models

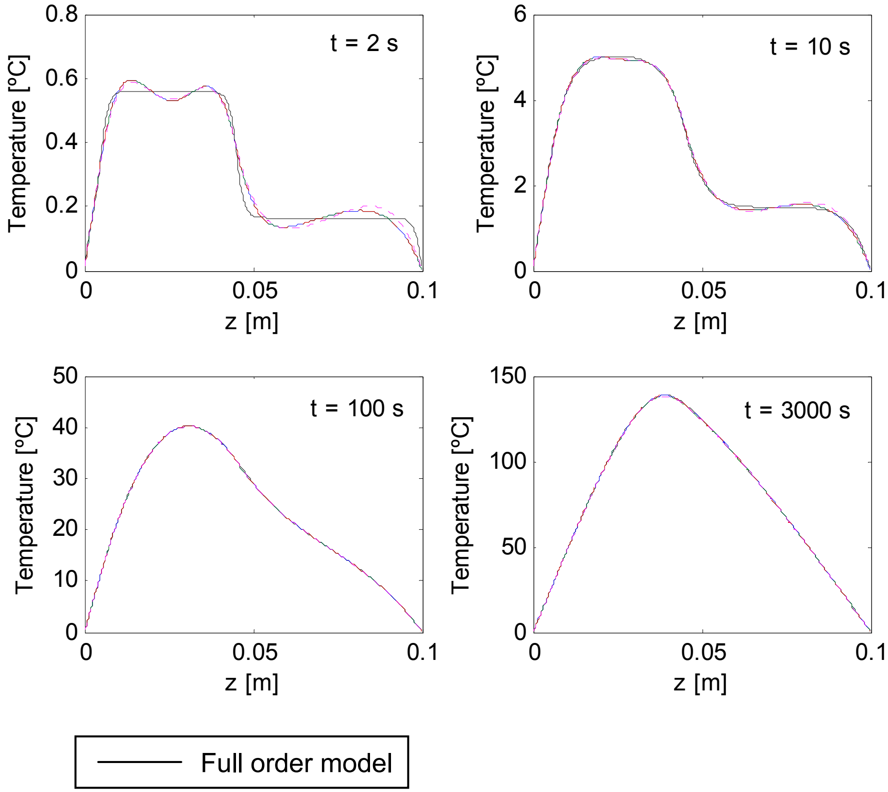
	G_d	G_s	ΔT_{max} [ºC]
POD model	0.97	2.02	0.423
Neural-POD model	9.66	8.153	0.689
MPE-POD model	2.08	3.75	0.695
P-POD model	5.61	5.00	0.423
P-POD-RS model	12.89	8.22	0.551

- G_d : Computational gain in the calculation of the derivatives.
- G_s : Computational gain in the simulation of the model.
- ΔT_{max} : Largest temperature deviation (error) of the POD model.

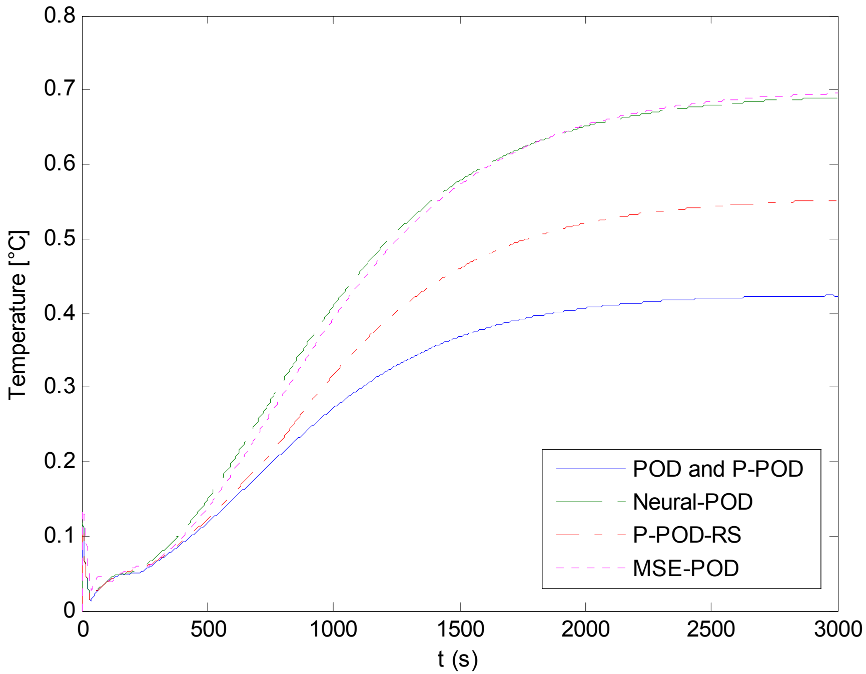
Some Simulation and Validation Results

Test 1: Steps of magnitude $1200 \cdot 10^3\ \mathrm{W \cdot m^{-3}}$ and $500 \cdot 10^3\ \mathrm{W \cdot m^{-3}}$ are applied to $u(t)$ and $d(t)$.

Temperature Profile of the bar



Maximum temperature deviation of the POD models with respect to the full order model



Conclusions and Future Research

Concluding Remarks

- In this thesis several POD-based MPC control schemes have been successfully designed for the reactor.

- Two approaches for reducing the number of state/output constraints of POD-based MPC Controllers have been presented:
  - Positive Polynomial Approach
  - Greedy Selection Algorithm

- Two alternative ways of accelerating the evaluation of nonlinear POD models have been proposed:
  - Neural Network Approach
  - Polynomial-POD model approach (the stability of these models has been discussed as well)

Future research

- Incorporate the nonlinear characteristics of the reactor into the POD-based MPC control schemes.
- Include into the reactor model, the dispersion/diffusion phenomena and the dynamic of the heat exchangers.
- Look for alternative methods for speeding up the evaluation of nonlinear POD models.
- Seek optimal ways of deriving POD basis vectors (SVD in tensors) for multidimensional systems.
- Establish guidelines to properly design the snapshot experiment.

Acknowledgments

- Promotor

  Bart De Moor

- Co-promotors

  Jairo Espinosa, Joos Vandewalle

- Positive Polynomials

  Moritz Diehl, Michel Baes

- Sequential feature selection methods

  Fabian Ojeda