



A Brief Perspective on Process Optimization

L. T. Biegler
Carnegie Mellon University
November, 2007

Annual AIChE Meeting
Salt Lake City, UT



Overview

Introduction

- Main Themes
- Modeling Levels

Closed Models

- PDEs and ROMs

NLPs and Model Integration

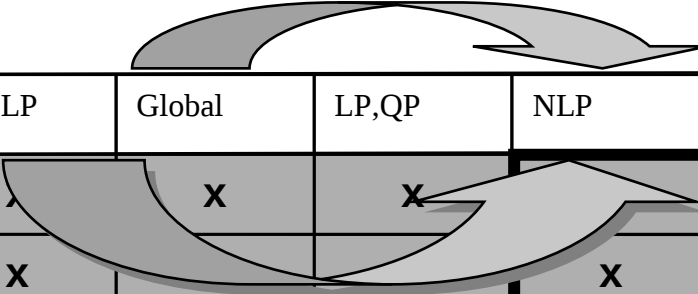
- LDPE Example

Fully Open Models

- Challenge Problem: NMPC

Conclusions

Optimization in Design, Logistics and Control



	MILP	MINLP	Global	LP, QP	NLP
HENS	X	X	X	X	
MENS	X	X			X
Separations	X	X			
Reactors		X	X	X	X
Equipment Design		X			X
Process Flowsheeting		X			X
Planning	X	X		X	
Scheduling	X	X		X	
Real-time optimization				X	X
Linear MPC				X	
Nonlinear MPC			X		X
Hybrid	X				X



Hierarchy for Optimization Models

1. Fully Open

- Completely declarative smooth models
- First, second derivatives cheap to calculate
- Solved with Newton-type methods

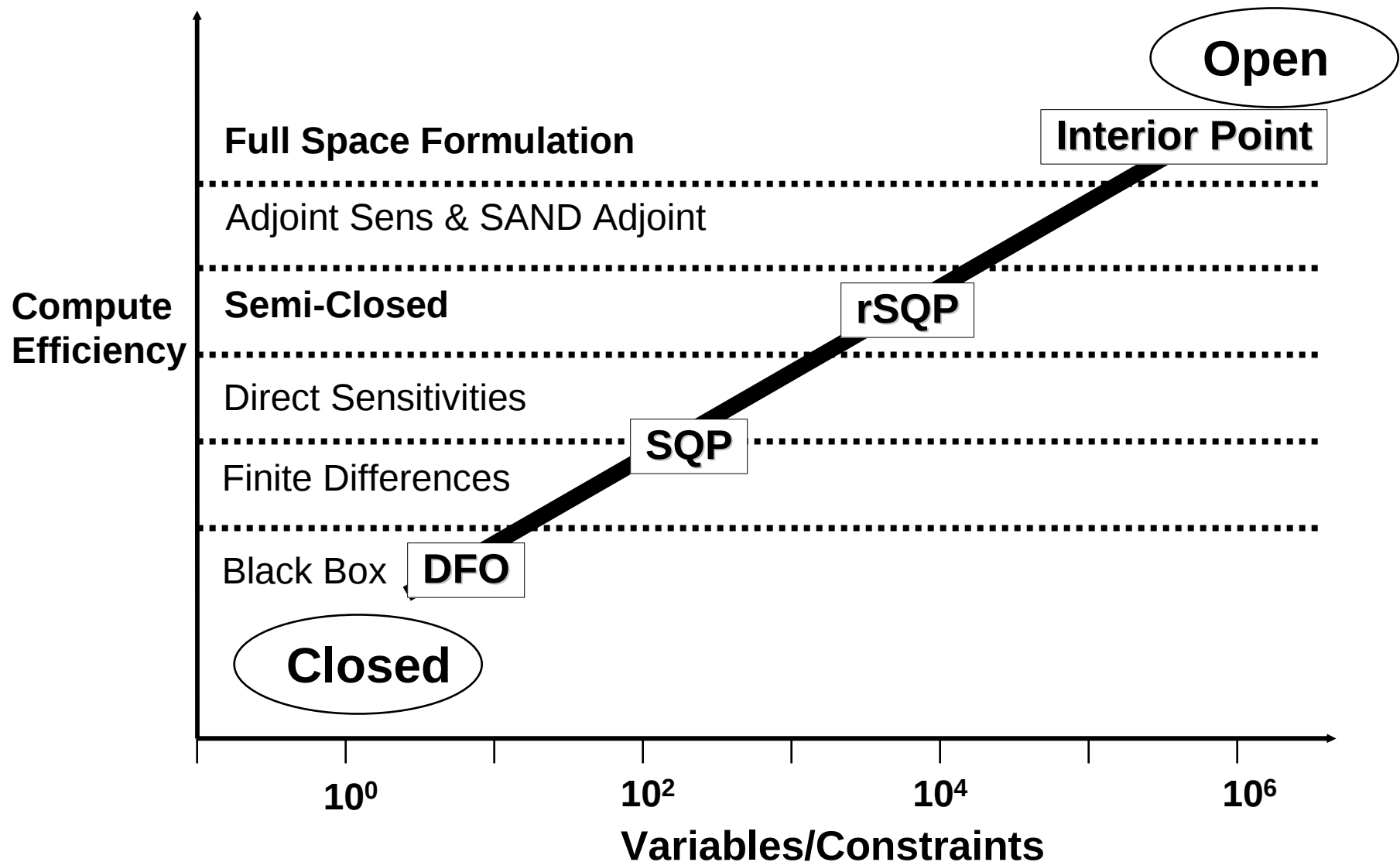
2. Semi-closed

- procedural models solved efficiently (e.g., DAE integrators)
- can be differentiated efficiently

3. Closed

- Procedural models, complex, special purpose solvers
- Smooth
- Difficult to differentiate

Hierarchy of Nonlinear Programming Formulations and Model Intrusion





Closed Optimization Models

Only function information available. Gradient and Hessian information cannot be obtained exactly

Nested iterative calculations in model (round-off error due to convergence tolerances)

How can these models be used for optimization?

Derivative Free Optimization

NOMAD, DFO, UOBYQA

Strong convergence theory, successful on difficult problems

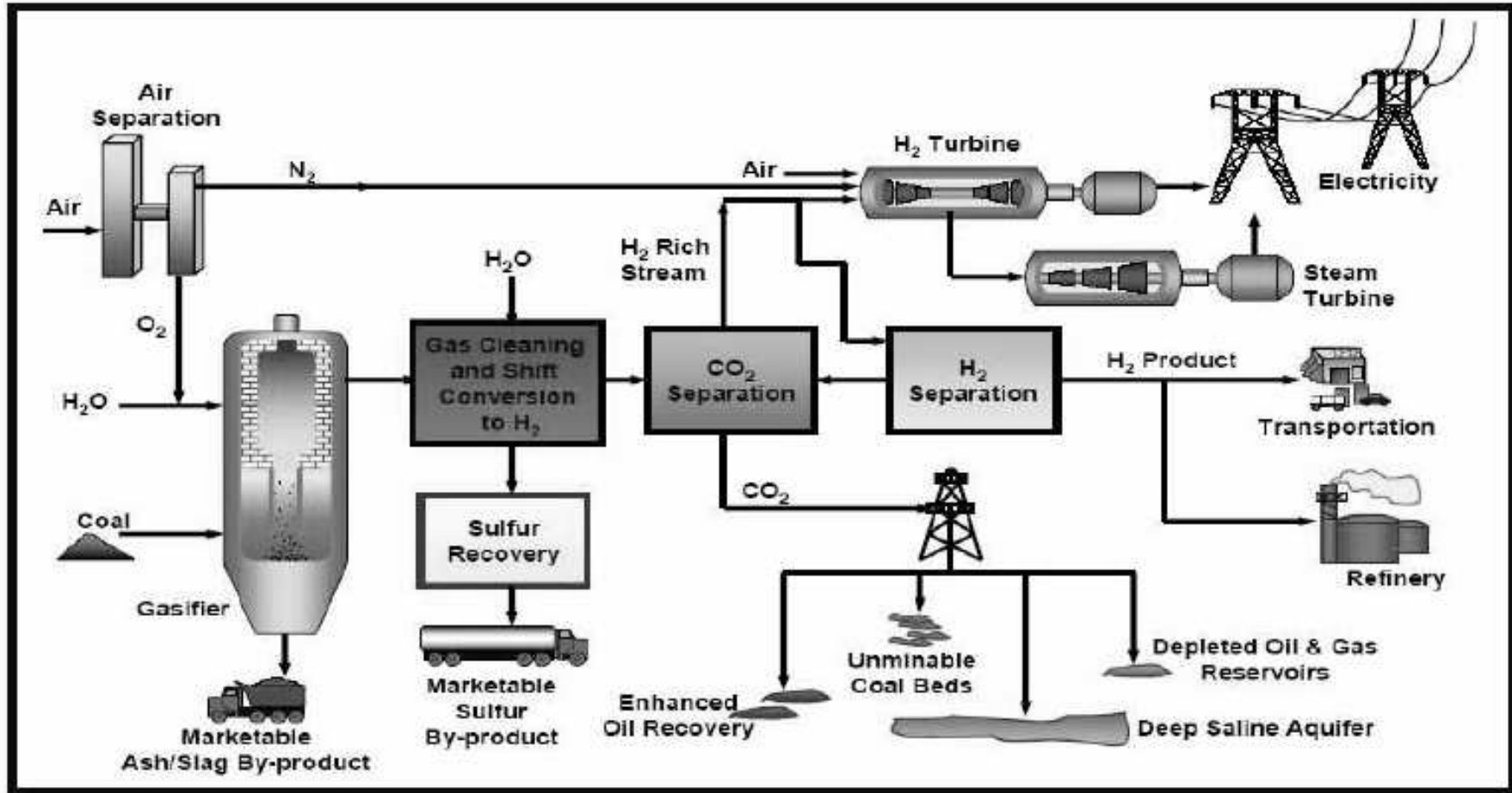
Reduced order models

convert to approximate models in analytic form
(manageable size)

Closed Detailed Models → Fully Open Simple Models

allows integration with models from other levels

Challenge for Process Optimization FutureGen Power Cycle (DOE/NETL)

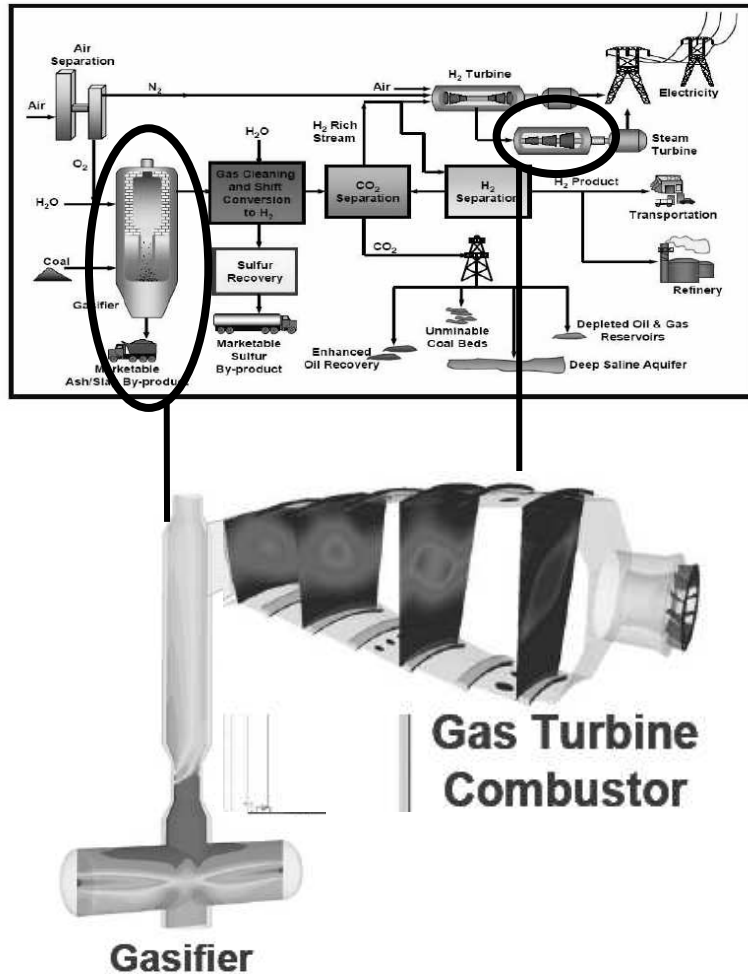


PDE/CFD: Gas Turbine, Gasifier

DAE: H_2/CO_2 Separation: PSA

Algebraic: Rest of Flowsheet

Advanced Process Engineering Cosimulation (Zitney and Syamlal, 2002)

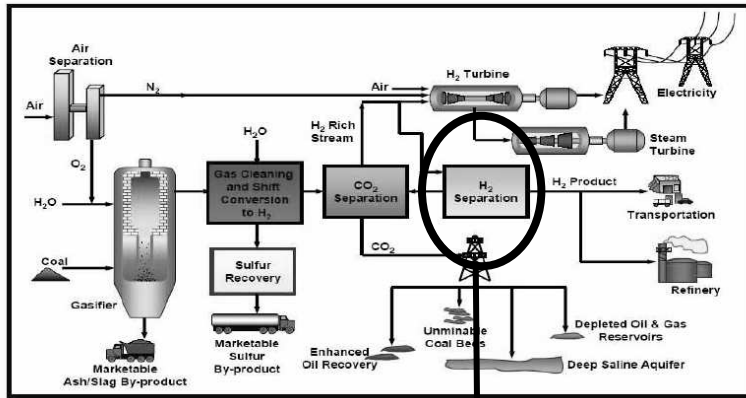


Detailed CFD Models

- Multiphase, reactive flow
- Specialized solution procedures
- Model requires ~1-10 CPU h
- Smooth responses expected but derivatives not available
- Few (<10) optimization variables

Goal: ROM for integration with other model equations

Optimal Operation for Gas Separation (Jiang, Fox, B., 2004)



Pressure Swing Adsorption in FutureGen Cycle

- Need for high purity H_2
- Possible technology to capture CO_2
- Respond quickly to changes in process demand
- Large, highly nonlinear dynamic separation
- Examples

• Large-scale H_2/CO Separation

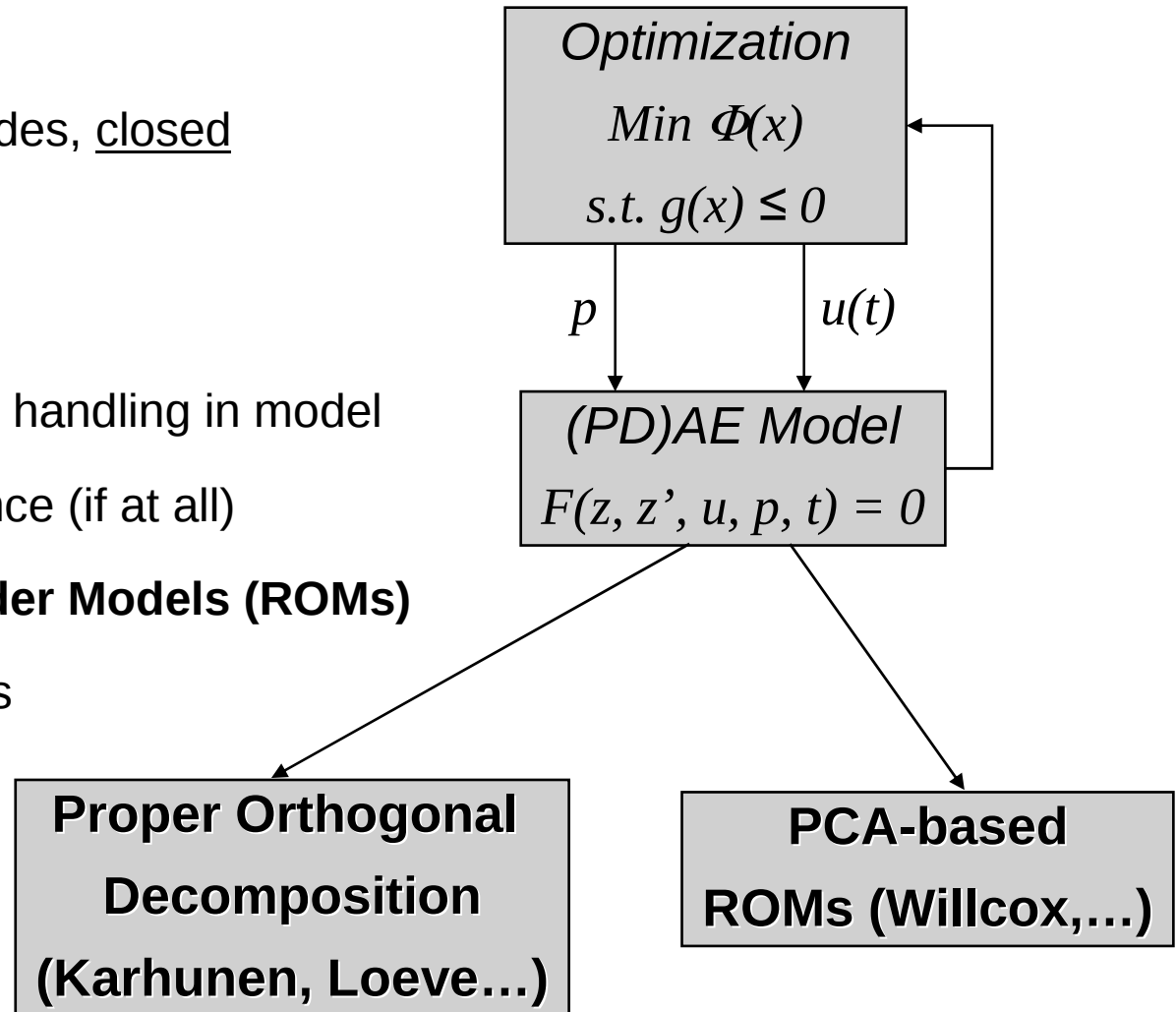
5 beds, 11 steps, $> 10^4$ variables
with dense Jacobians, 10^4 DAEs

Simulated Moving Bed Separation

6 columns, 4 zones, > 100
variables/ DAEs with dense
Jacobians

Black Box Optimization (Closed Models)

- + Easy to interface
- + Easily applied to legacy codes, closed models
- Repeated model solutions
- Failure prone, no constraint handling in model
- Derivatives by finite difference (if at all)
- **(Fully Open) Reduced Order Models (ROMs)** for optimization formulations



PSA Optimization

4 step Methane/Hydrogen Separation

Variable	Value
High operating bed pressure (P_H)	5×10^5 bars
Low operating bed pressure (P_L)	1.5×10^5 bars
Pressurization step time (t_p)	5 sec
Adsorption step time (t_a)	50 sec

$Max \quad Recovery_{H_2}$

$s.t.$ ROM Equations
 $4.8 \times 10^5 \leq P_H \leq 5.2 \times 10^5$
 $1.3 \times 10^5 \leq P_L \leq 1.7 \times 10^5$
 $3 \leq t_p \leq 7$
 $47 \leq t_a \leq 53$
 $Purity_{H_2} \geq 0.998$

PSA Model: 3200 DAEs (several CPU hrs to optimize)
POD Model Reduction: reduced to just 200 DAEs

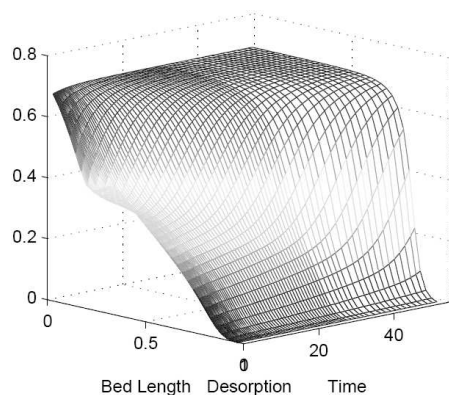
**DAEs in ROM were converted to algebraic equations
by temporal discretization (simultaneous approach)**

**The resultant non-linear program (NLP) was solved using
IPOPT solver in AMPL: optimum obtained in 3 CPU min**

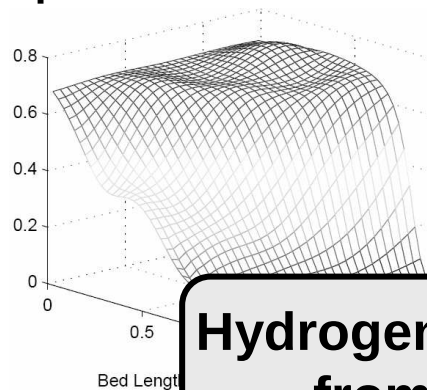
PSA/ROM Optimization Results

Profiles for gas-phase methane mole fraction

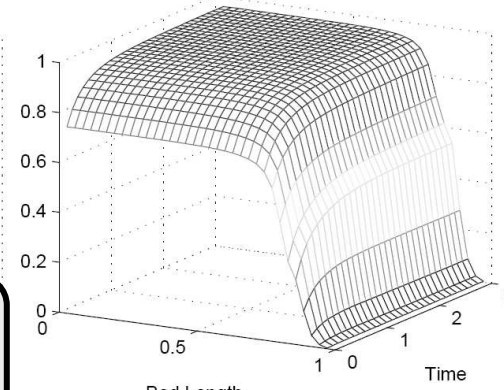
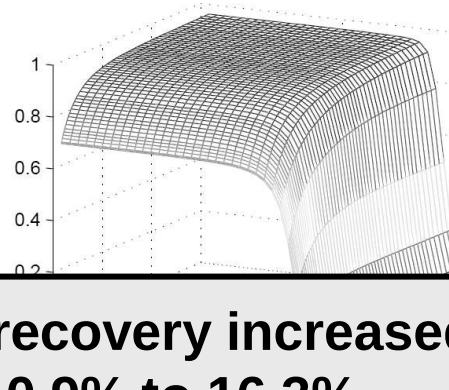
Adsorption



Rigorous Model



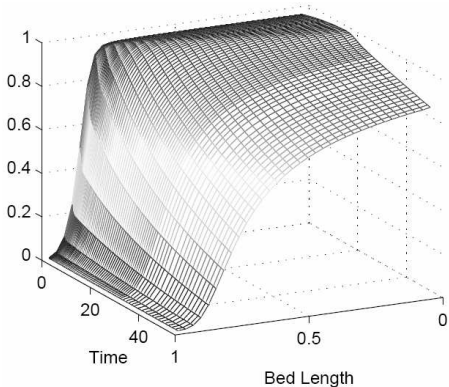
Depressurization



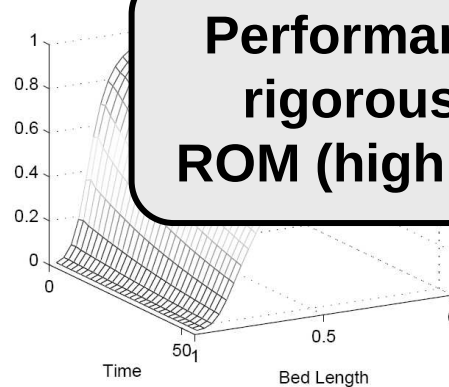
ROM

**Hydrogen recovery increased
from 10.9% to 16.3%**

Desorption

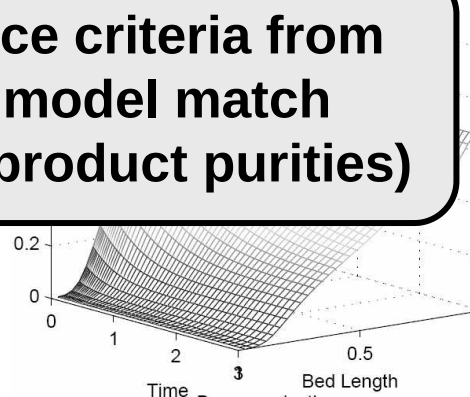


Rigorous Model

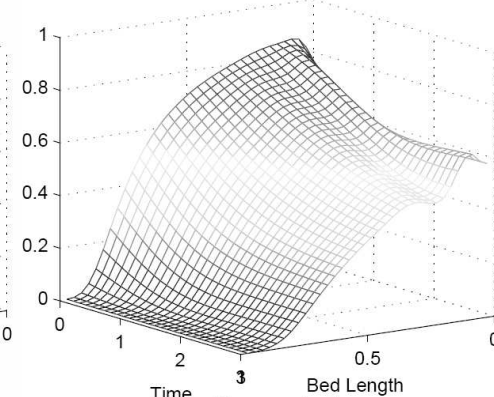


ROM

Pressurization



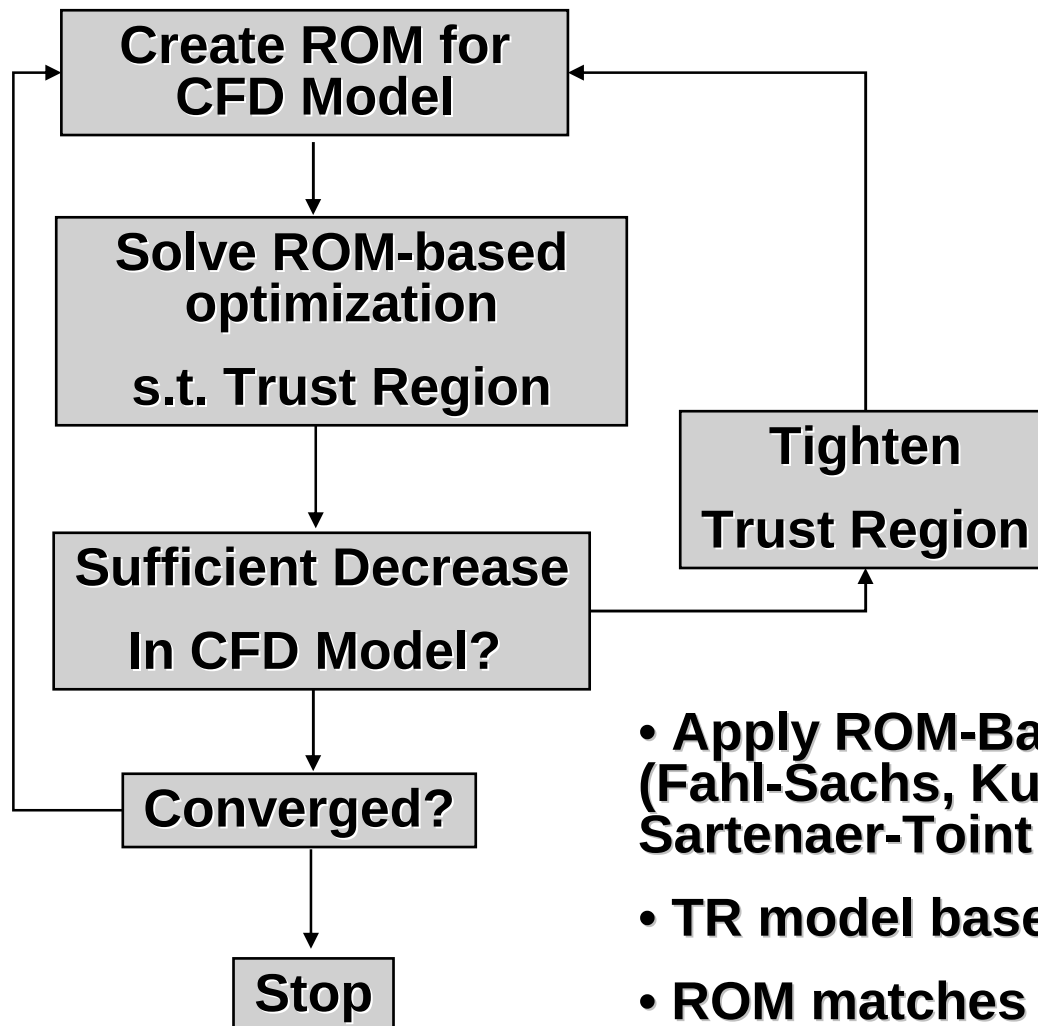
Rigorous Model



ROM

**Performance criteria from
rigorous model match
ROM (high product purities)**

ROM Sufficiently Accurate for Optimization?



- Apply ROM-Based Trust Region Strategy (Fahl-Sachs, Kunisch-Volkwein, Gratton-Sartenaer-Toint...)
- TR model based on ROM optimization
- ROM matches functions/gradients



Fully Open Optimization Models

- Exact first (and higher!) derivatives available
- Single model solution – no nested calculations
- Basis for extending NLP strategies to MINLP, global optimization
- Which NLP algorithms should be used?
 - Fast convergence properties
 - Low complexity in dealing with extending to large-scale systems
 - Exploit structures of model and optimization formulations
- Essential for many time-critical optimizations (RTO, NMPC, DRTO)

NLP

Min $\Phi(x)$

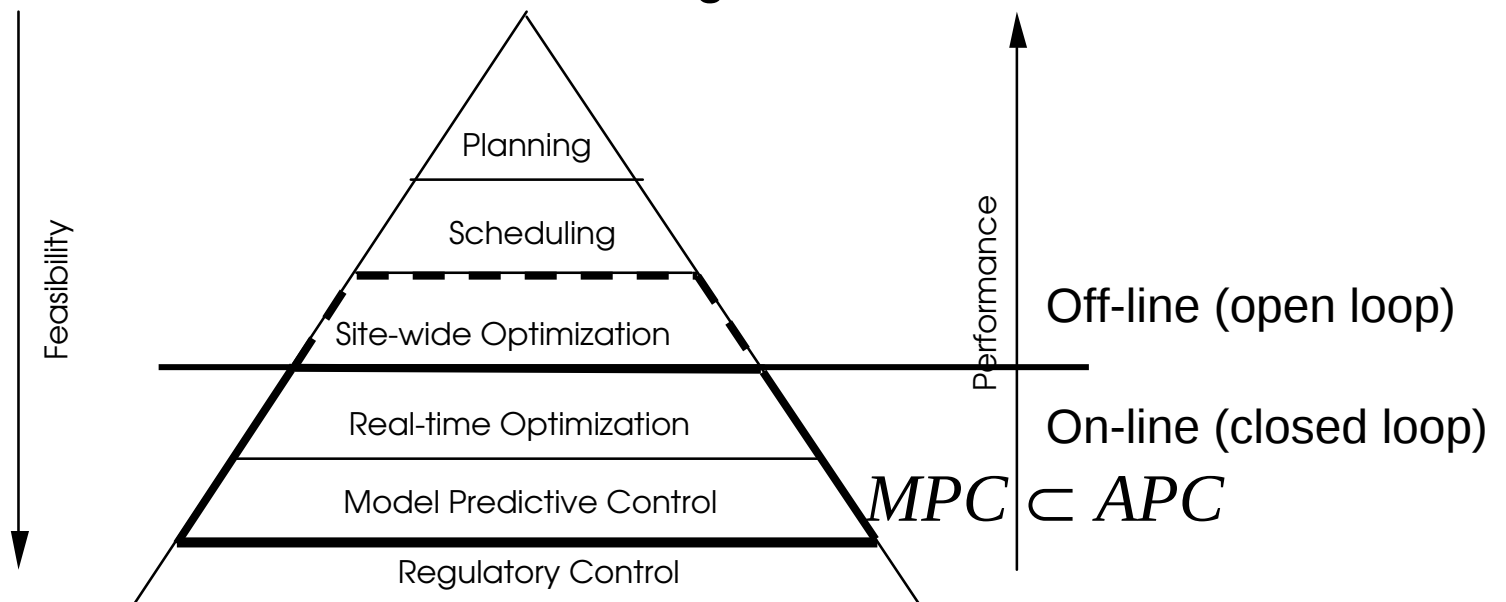
s.t. $c(x) = 0$

$x_L \leq x \leq x_U$

Decision Pyramid for Process Operations

Real-time Optimization and Advanced Process Control

- Fewer discrete decisions
- Many nonlinearities
- Frequent, “on-line” time-critical solutions
- Higher level decisions must be feasible
- Performance communicated for higher level decisions





Dynamic Real-time Optimization

Integrate On-line Optimization/Control with Off-line Planning

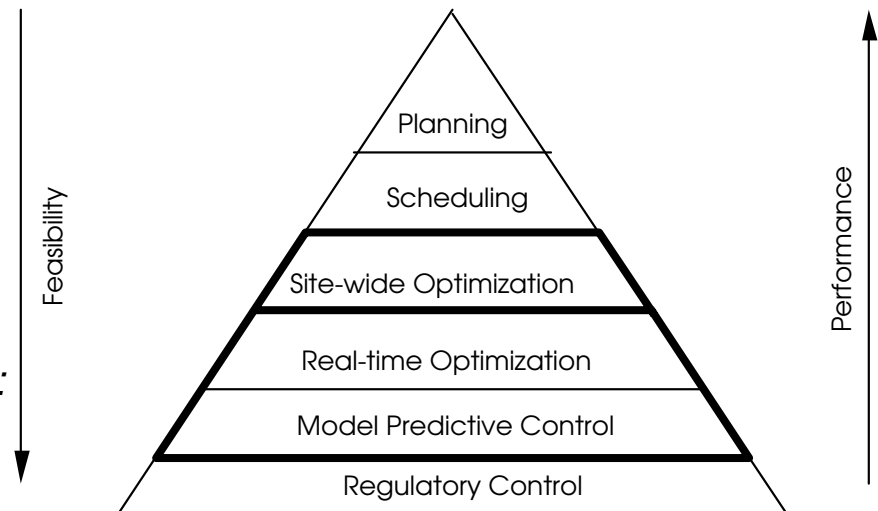
- Consistent, first-principle models
- Consistent, long-range, multi-stage planning
- Increase in computational complexity
- Time-critical calculations

Applications

- Batch processes
- Grade transitions
- Cyclic reactors (coking, regeneration...)
- Cyclic processes (PSA, SMB...)

Continuous processes are never in steady state:

- Feed changes
- Nonstandard operations
- Optimal disturbance rejections



Simulation environments and first principle dynamic models are widely used for off-line studies

Can these results be implemented directly on-line for large-scale systems?



Dynamic Optimization Engines

Evolution of NLP Solvers:

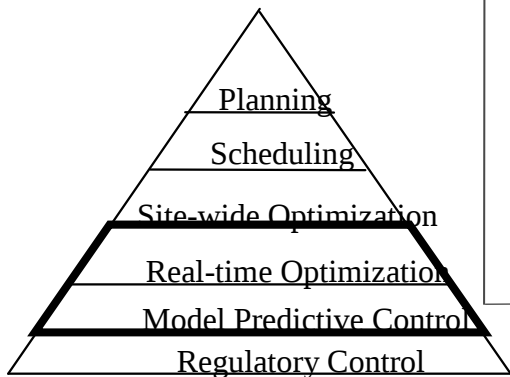
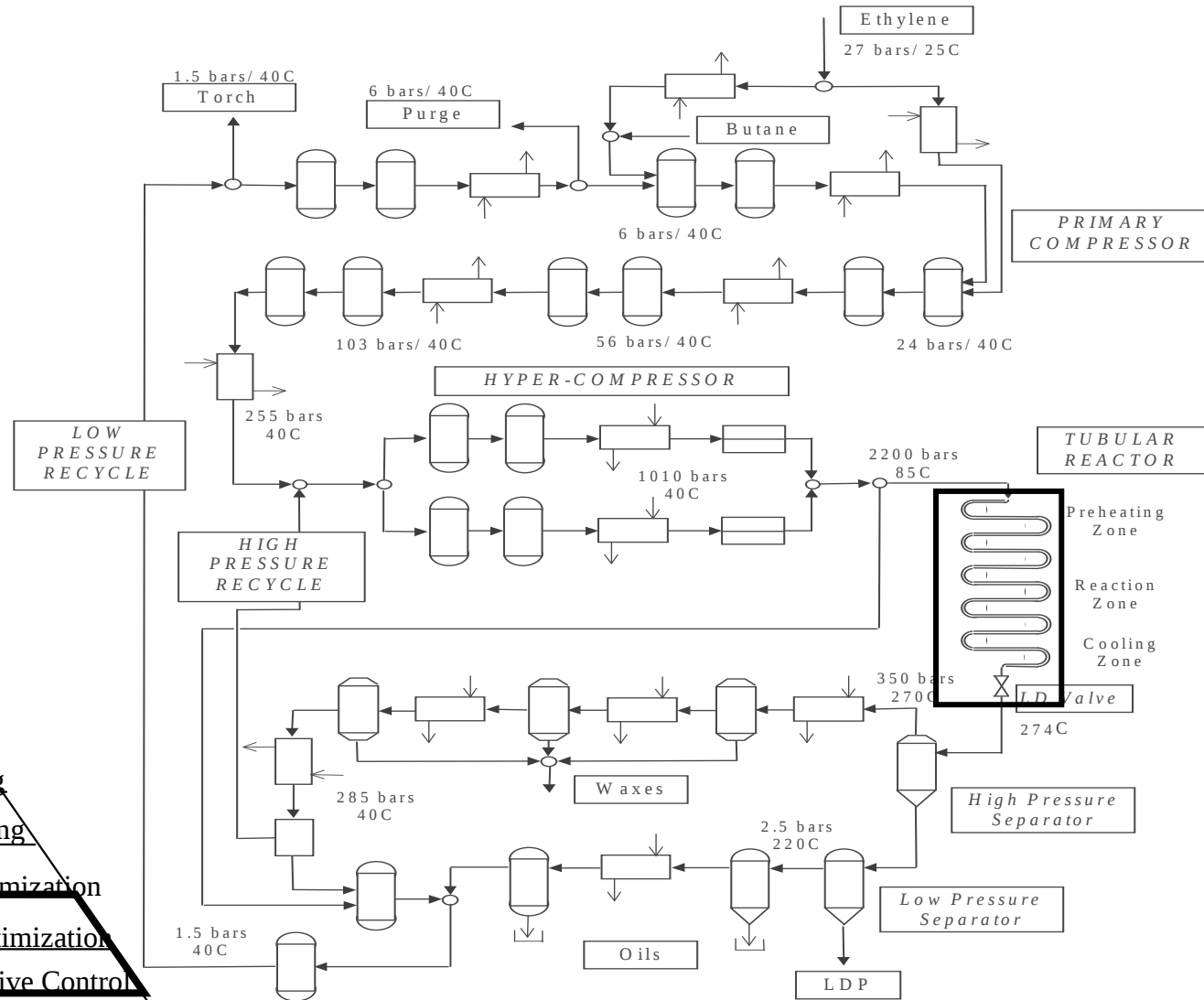
→ *for dynamic optimization, control and estimation*

SQP → rSQP → Full-space
Barrier

E.g., ***IPOPT*** - Simultaneous dynamic optimization
over 1 000 000 variables and constraints

Object Oriented Codes tailored to structure, sparse linear
algebra and computer architecture (e.g., IPOPT 3.3)

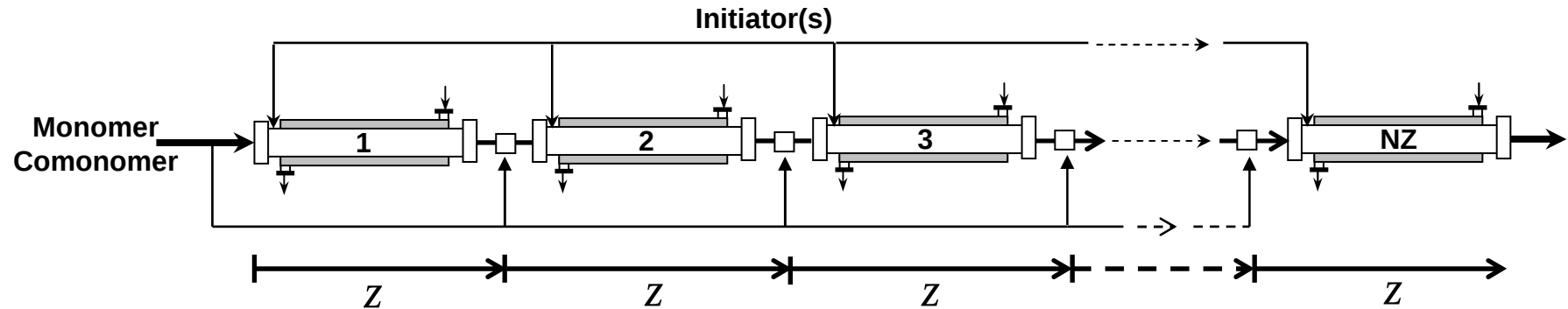
LDPE Plant- Multistage Dynamic Optimization



High pressure reaction (>2000 atm) through mile-long reactor coil

Highly nonlinear model with 1000's of chemical species (moment models)

Large-Scale Parameter Estimation



Material & Energy

Physical Properties

Zone Transitions

$$\left\{ \begin{array}{l} F_j \left[\frac{dy_j(z)}{dz}, y_j(z), w_j(z), z, \pi_j, \Pi \right] = 0 \\ G_j \left[y_j(z), w_j(z), z, \pi_j, \Pi \right] = 0 \\ y_j(0) = \phi(y_{j-1}(z_{L_{j-1}}), F_{f_j}) \end{array} \right\} j \in \{1..NZ\}$$

500 ODEs

1000 AEs

× Stiffness + Highly Nonlinear + Parametric Sensitivity + Algebraic Coupling

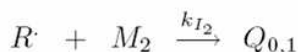
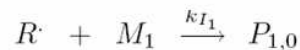
Large-Scale Parameter Estimation

□ Complex Kinetic Mechanisms

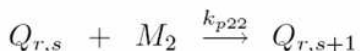
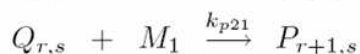
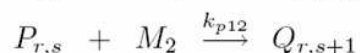
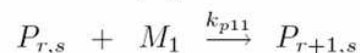
Initiator decomposition



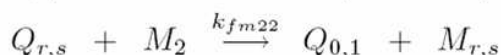
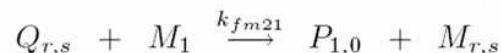
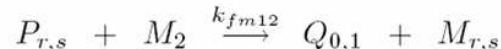
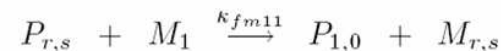
Chain Initiation



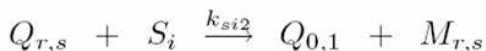
Chain Propagation



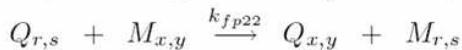
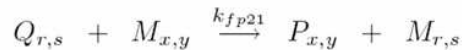
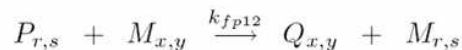
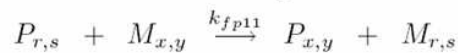
Chain Transfer to Monomer



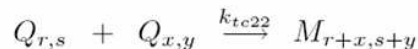
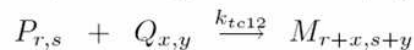
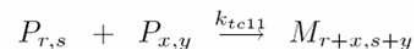
Chain Transfer to Solvent



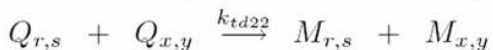
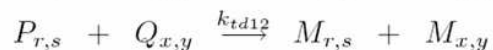
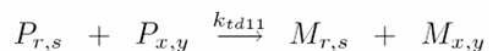
Chain Transfer to Polymer



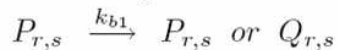
Termination by Combination



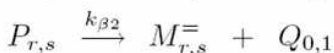
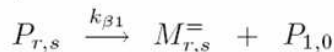
Termination by Disproportionation



Backbiting



β-scission



$$k = k_0 \exp\left(-\frac{E_a + P E_v}{RT}\right)$$

~ 35 Elementary Reactions
~100 Kinetic Parameters



Large-Scale Parameter Estimation

- Parameter Estimation for Industrial Applications
 - Use Rigorous Model to Match Plant Data Directly
 - Start with Standard Least-Squares Formulation

$$\begin{aligned}
 \min_{\Pi, \pi_{k,j}} & \sum_{k=1}^{NS} \sum_{j=1}^{NZ} \sum_{i=1}^{NM(j)} \left(y_{k,j}(z_i) - y_{k,j,i}^M \right)^T \mathbf{V}_y^{-1} \left(y_{k,j}(z_i) - y_{k,j,i}^M \right) \\
 & + \sum_{k=1}^{NS} \left(w_{k,NZ} - w_{k,NZ}^M \right)^T \mathbf{V}_w^{-1} \left(w_{k,NZ} - w_{k,NZ}^M \right) \\
 s.t. & \\
 & \mathbf{F}_{k,j} \left[\frac{dy_{k,j}(z)}{dz}, y_{k,j}(z), w_{k,j}(z), z, \pi_{k,j}, \Pi \right] = 0 \\
 & \mathbf{G}_{k,j} \left[y_{k,j}(z), w_{k,j}(z), z, \pi_{k,j}, \Pi \right] = 0 \\
 & y_{k,j}(0) = \phi(y_{k,j-1}(z_{L_{k,j-1}}), F_{f_{k,j}}) \\
 & j \in \{1..NZ\}, \quad k \in \{1..NS\}
 \end{aligned}$$

} Least-Squares
 } Rigorous Reactor Model

- Special Case of Multi-Stage Dynamic Optimization Problem
 - Solve using Simultaneous Collocation-Based Approach

1 data set 500 ODEs 1000 AEs	x 6 	6 data sets 3000 ODEs 6000 AEs
---	----------------	---

Multi-zone Reactor Parameter Estimation

- **Data Sets:** Operating Conditions and Properties for Different Grades
- **Match:** Temperature Profiles and Product Properties
 - **On-line Adjusting Parameters** → Track Evolution of Disturbances
 - **Kinetic Parameters** → Apply to Rigorous Models
 - **Add Data for Process Inputs (EVM)** → remove additional uncertainties
- **Single Data Set** (On-line Adjusting Parameters)

Grade	Constraints	Parameters	LB	UB	Iterations	CPUs	NZJ	NZH
A	11955	32	374	361	11	17.03	166425	87954
B	11283	32	374	361	8	10.06	138666	76890

- **Multiple Data Sets** (On-line Adjusting Parameters + Kinetics)

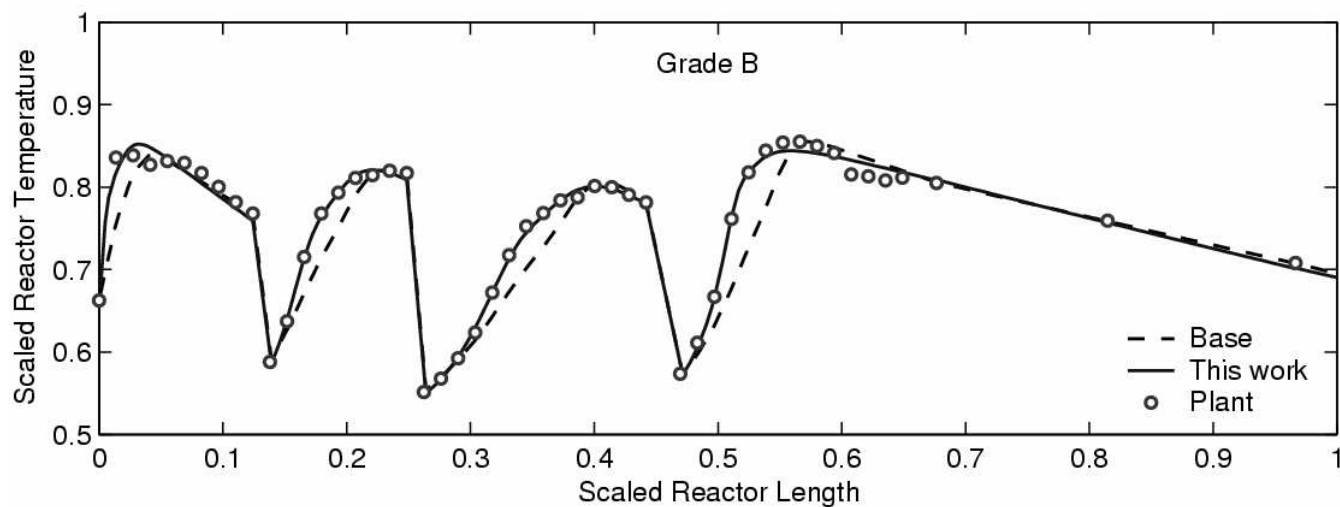
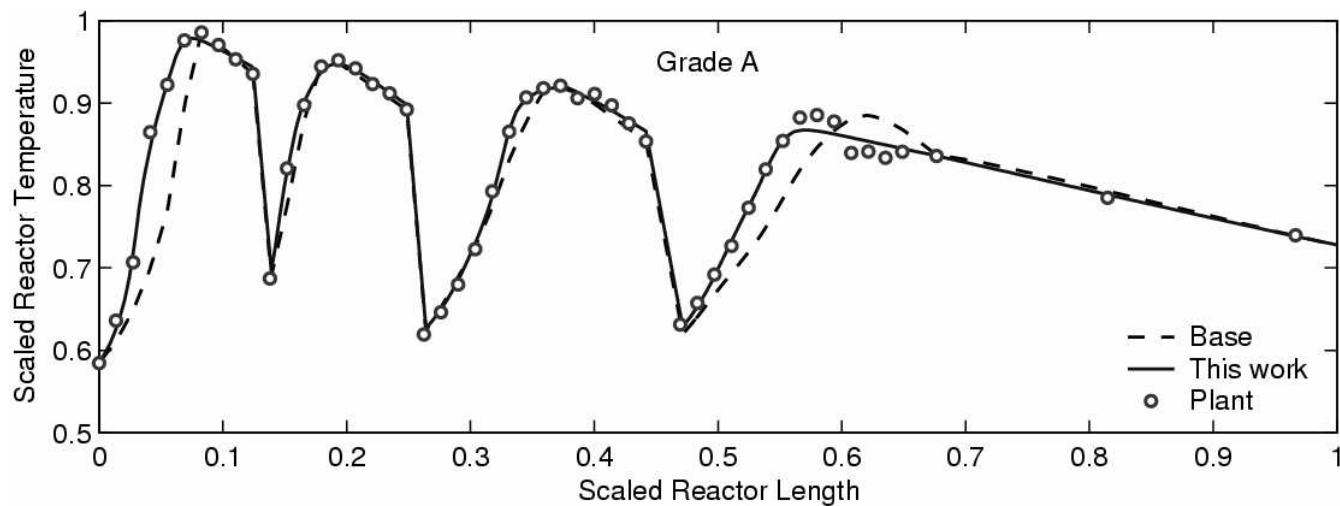
Data Sets	Constraints	DOF	LB	UB	Iterations	CPUs	NZJ	NZH
3	33900	121	1246	1207	68	451.51	520275	552738
6	68421	217	2467	2389	58	900.21	1058412	1119258

- **EVM** (On-line Adjusting Parameters + Kinetics + Process Inputs)

Data Sets	Constraints	DOF	LB	UB	Iterations	CPUs	NZJ	NZH
6 (EVM)	68627	529	2653	2575	71	1010.74	1059512	1119780
6 (SLS)	68421	217	2467	2389	58	900.21	1058412	1119258

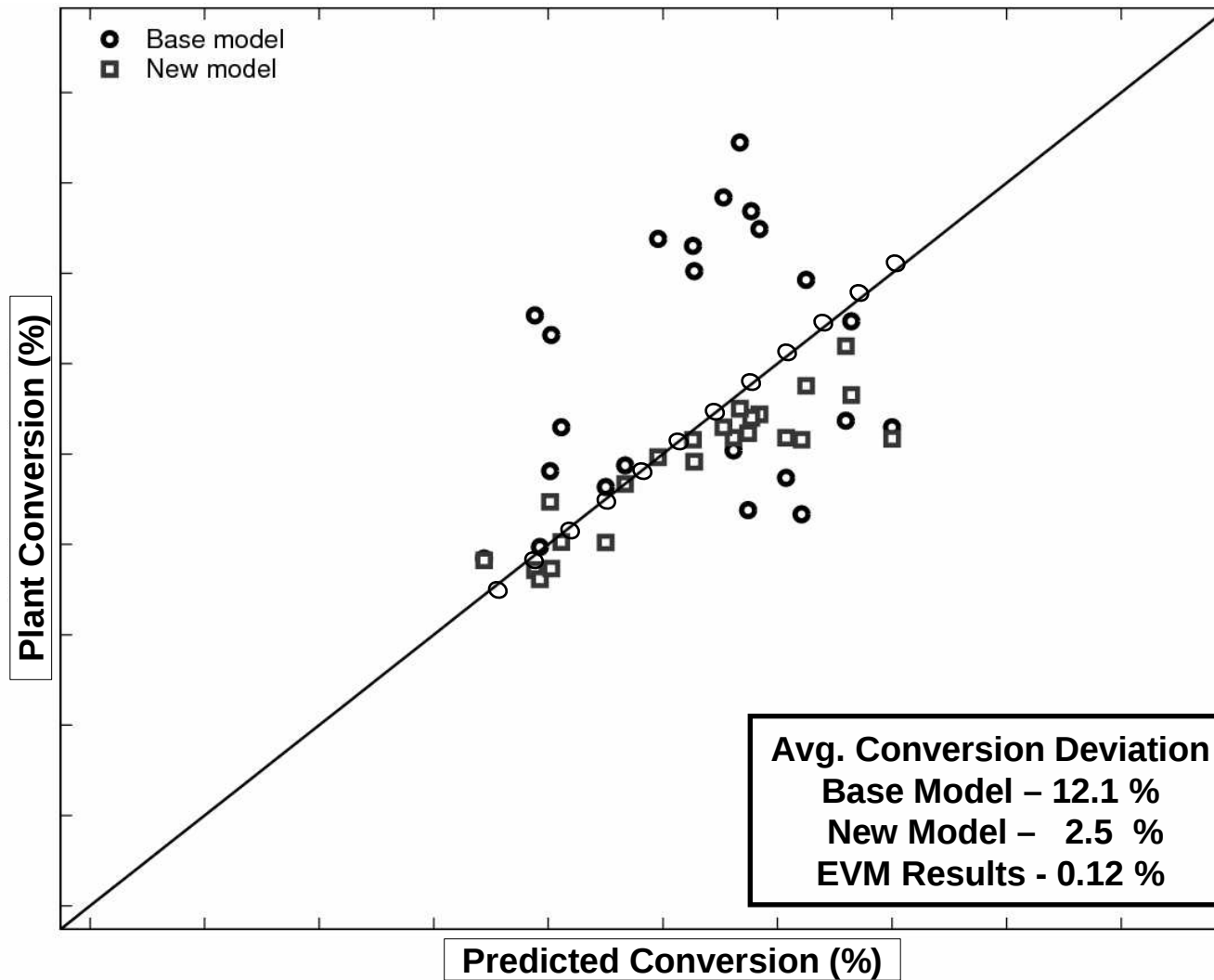
Large-Scale Parameter Estimation

Improved Match of Reactor Temperatures Profile

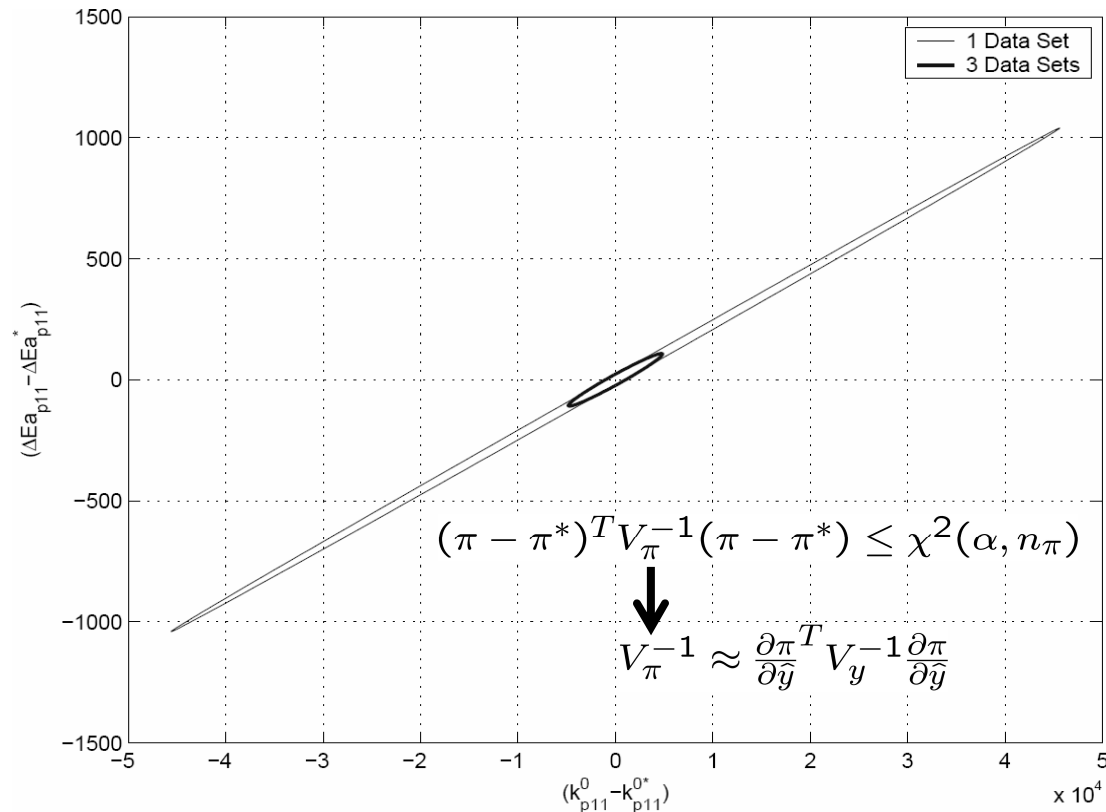


Industrial Case Study

□ Results - Reactor Overall Monomer Conversion (up to 20 Different Grades)



IPOPT Factorization Byproduct



KKT matrix factored by indefinite symmetric factorization

- Without regularization at solution → sufficient second order conditions and uniquely estimated parameters
- Reduced Hessian cheaply available to calculate confidence regions

Parameter Estimation in Parallel Architectures

Exploit Structure of KKT Matrix – Laird, B. 2006

$$\min_{\Pi, \pi_{k,j}} \sum_{k=1}^{NS} \sum_{j=1}^{NZ} \sum_{i=1}^{NM(j)} (y_{k,j}(z_i) - y_{k,j,i}^M)^T V_y^{-1} (y_{k,j}(z_i) - y_{k,j,i}^M) + \sum_{k=1}^{NS} (w_{k,NZ} - w_{k,NZ}^M)^T V_w^{-1} (w_{k,NZ} - w_{k,NZ}^M)$$

s.t.

$$\begin{aligned} F_{k,j} \left[\frac{dy_{k,j}(z)}{dz}, y_{k,j}(z), w_{k,j}(z), z, \pi_{k,j}, \Pi \right] &= 0 \\ G_{k,j} [y_{k,j}(z), w_{k,j}(z), z, \pi_{k,j}, \Pi] &= 0 \\ y_{k,j}(0) &= \phi(y_{k,j-1}(z_{L_{k,j-1}}), F_{f_{k,j}}) \\ j &\in \{1..NZ\}, \quad k \in \{1..NS\} \end{aligned}$$

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned}$$

$$\begin{bmatrix} H_k & A_k & -I \\ A_k^T & 0 & 0 \\ V_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \nu \end{bmatrix} = - \begin{bmatrix} r_x \\ r_\lambda \\ X_k V_k e - \mu_\ell e \end{bmatrix}$$

Direct Factorization MA27
Memory Bottlenecks
Factorization Time Scales
Superlinearly with Data sets

$$\begin{aligned} \min \quad & \sum_{k=1}^{NS} f_k(x_k) \\ \text{s.t.} \quad & c_k(x_k, \Pi) = 0 \\ & x_k, \Pi \geq 0 \end{aligned}$$

$$\begin{bmatrix} \boxed{K_1} & & & & Q_1 \\ & \boxed{K_2} & & & Q_2 \\ & & \dots & & \vdots \\ & & & \boxed{K_N} & Q_N \\ \hline Q_1^T & Q_2^T & \dots & Q_N^T & D_\Pi \end{bmatrix} \begin{bmatrix} \Delta s_1 \\ \Delta s_2 \\ \vdots \\ \Delta s_N \\ \Delta \Pi \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \\ r_\Pi \end{bmatrix}$$

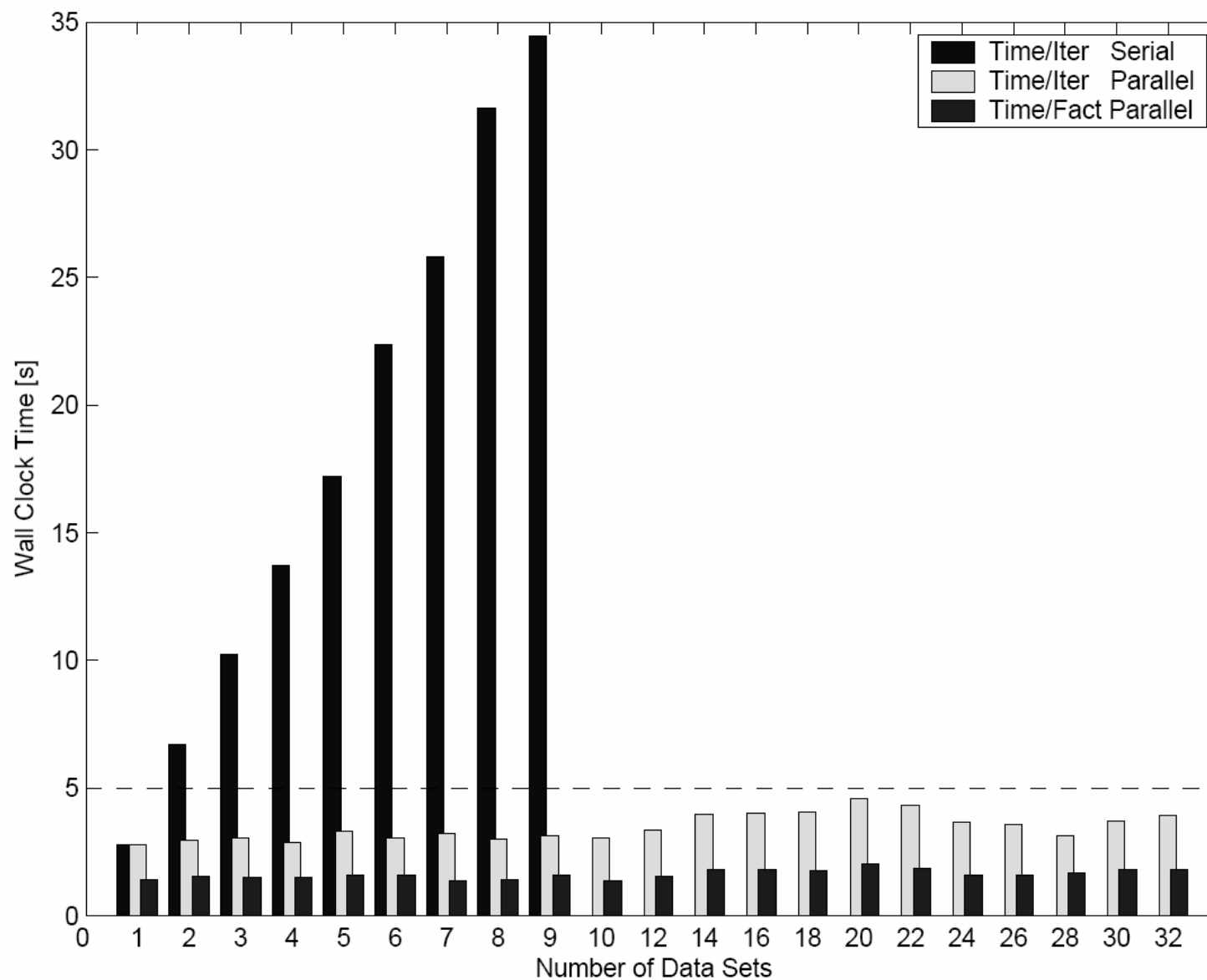
Block-bordered Diagonal Structure
Coarse-Grained Parallelization using
Schur Complement Decomposition

IPOPT 3.x architecture supports tailored structured decompositions



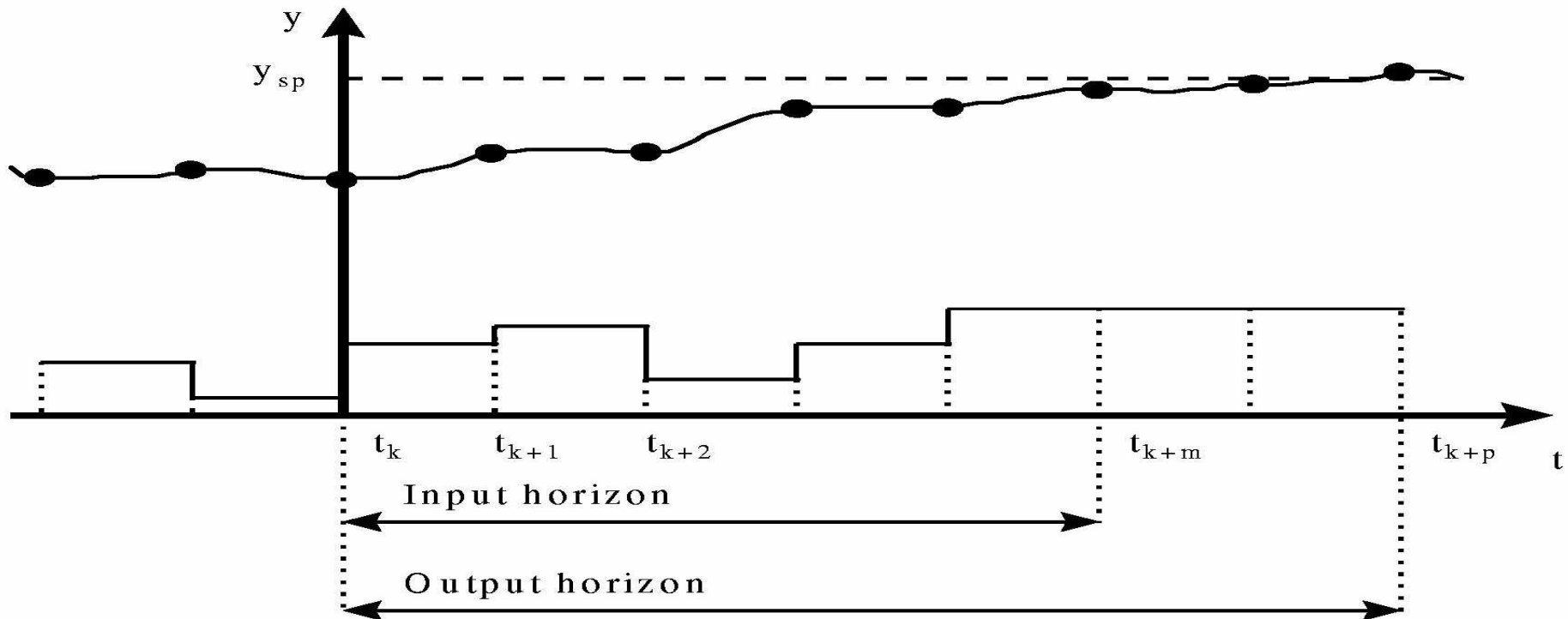
Parameter Estimation in Parallel Architectures

Computational Results – LDPE Reactor EVM Problem



On-line Issues: Model Predictive Control (NMPC)

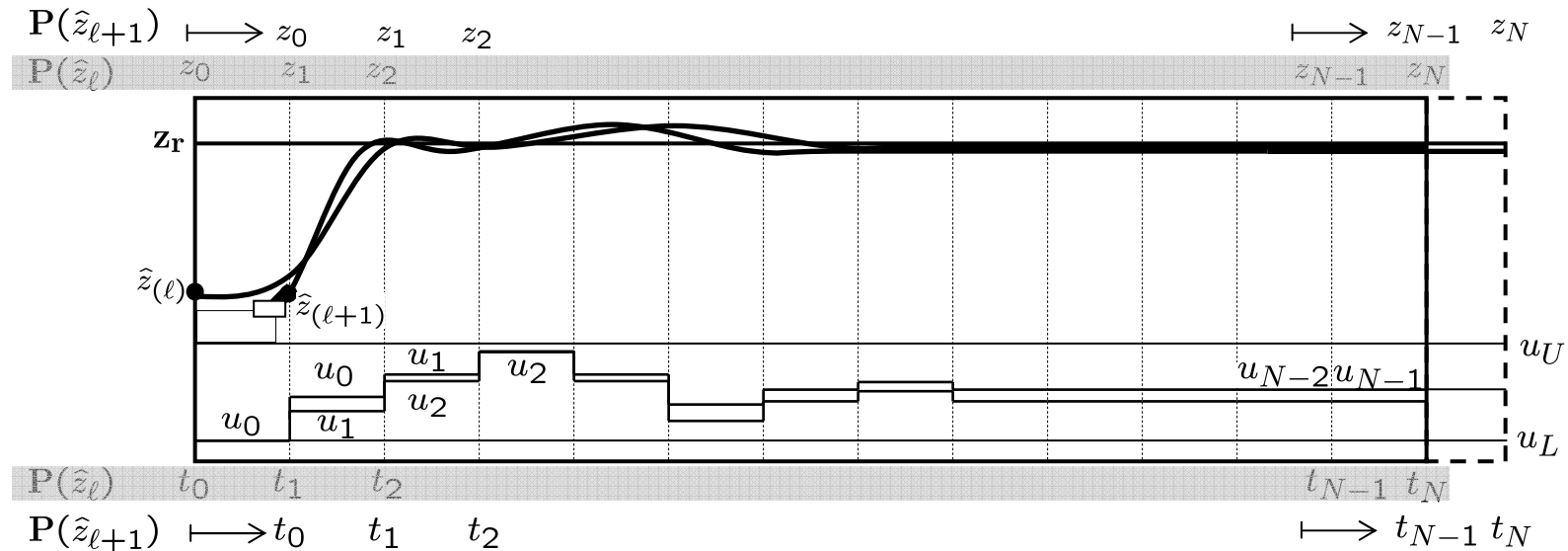
Zavala, Laird, B. (2006, 2007)



$$\begin{aligned}
 \min_u \quad & J(x(k)) = \sum_{l=0}^N \varphi(z_l, u_l) + E(z_N) \\
 \text{s.t.} \quad & z_{l+1} = f(z_l, u_l) \\
 & z_0 = x(k) \\
 & \text{Bounds}
 \end{aligned}$$

Challenge: Computational Delay → Performance and Stability?

Nonlinear Model Predictive Control – Parametric Problem

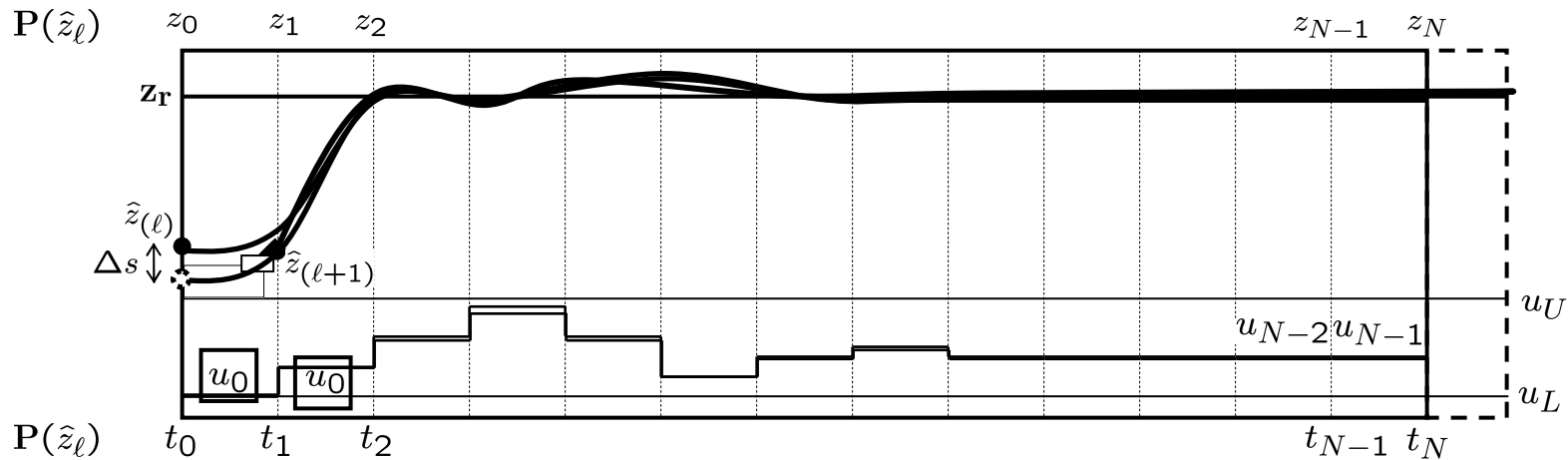


$$\begin{aligned}
 \mathcal{P}(x(k), N) \quad & \min_{v_{l|k}} J(x(k), N) = F(z_{k+N|k}) + \sum_{l=k}^{k+N-1} \psi(z_{l|k}, v_{l|k}) \\
 \text{s. t.:} \quad & z_{l+1|k} = f(z_{l|k}, v_{l|k}), \quad l = k, \dots, k+N-1 \\
 & \boxed{z_{k|k} = x(k) = p_0} \\
 & z_{l|k} \in \mathbb{X}, \quad z_{k+N|k} \in \mathbb{X}_f, \quad v_{l|k} \in \mathbb{U}.
 \end{aligned}
 \quad \left. \vphantom{\begin{aligned} \mathcal{P}(x(k), N) } \right\} \begin{array}{l} \mathcal{P}(p_0) \\ \mathcal{P}(p) \end{array}$$

$$\begin{aligned}
 \mathcal{P}(x(k+1), N) \quad & \min_{v_{l|k+1}} J(x(k+1), N) = F(z_{k+N+1|k+1}) + \sum_{l=k+1}^{k+N} \psi(z_{l|k+1}, v_{l|k+1}) \\
 \text{s. t.:} \quad & z_{l+1|k+1} = f(z_{l|k+1}, v_{l|k+1}), \quad l = k+1, \dots, k+N \\
 & \boxed{z_{k|k+1} = x(k+1) = p} \\
 & z_{l|k+1} \in \mathbb{X}, \quad z_{k+N+1|k+1} \in \mathbb{X}_f, \quad v_{l|k} \in \mathbb{U}.
 \end{aligned}$$

Advanced Step NMPC

Combine advanced step with sensitivity to solve NLP in background
(between steps) – not on-line



$$\begin{aligned}
 \min_{v_{l|k}} \quad & J(x(k), N) = F(z_{k+N|k}) + \sum_{l=k}^{k+N-1} \psi(z_{l|k}, v_{l|k}) \\
 \text{s. t.} \quad & z_{k+1|k} = f(x(k), u(k)), \\
 & z_{l+1|k} = f(z_{l|k}, v_{l|k}), \quad l = k+1, \dots, k+N-1 \\
 & z_{k|k} = z_0
 \end{aligned}$$

$$\begin{bmatrix} W_k & A_k & -I \\ A_k^T & 0 & 0 \\ V_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta v \end{bmatrix} = K \Delta v$$

$$\begin{bmatrix} K & E_0 \\ E_1^T & 0 \end{bmatrix} \begin{bmatrix} \Delta v \\ \Delta s \end{bmatrix} = - \begin{bmatrix} 0 \\ \hat{z}(\ell+1) - z_1^* \end{bmatrix}$$

$$z_{l|k} \in \mathbb{X}, \quad z_{k+N|k} \in \mathbb{X}_f, \quad v_{l|k} \in \mathbb{U}.$$

Solve $P(z_\ell)$ in background (between t_0 and t_1)

Sensitivity to updated problem to get (z_0, u_0)

Solve $P(z_{\ell+1})$ in background with new (z_0, u_0)

Industrial Case Study – Grade Transition Control

Process Model: 289 ODEs, 100 AEs

$$\min \int_{t_\ell}^{t_\ell+t_p} (w_{C_4}(t) - w_{C_4}^r)^2 + (F_{C_4}(t) - F_{C_4}^r)^2 + (F_{pu}(t) - F_{pu}^r)^2$$

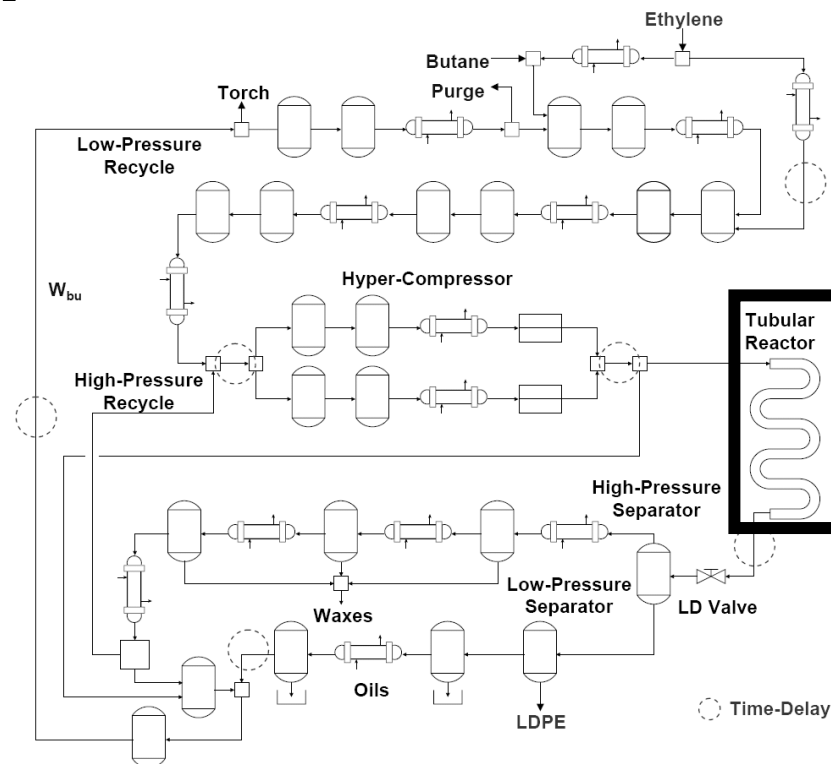
s.t.

$$\begin{aligned} & \text{PDEs+ODEs} \\ & z(t = t_\ell) = \hat{z}(\ell) \\ & z_L \leq z \leq z_U \\ & y_L \leq y \leq y_U \\ & u_L \leq u \leq u_U \end{aligned}$$

Simultaneous Collocation-Based Approach

$$\begin{aligned} \min & \sum_{m=0}^{M-1} \left\{ \sum_{k=0}^{K_m-1} \varphi_m^k(z_m^k, y_m^k, u_m^k, \pi_m, \Pi) \right\} \\ \text{s.t.} & \\ & z_m^{k+1} = z_m^k + \mathbf{A}_m y_m^k \\ & 0 = h_{m,k}(z_m^k, y_m^k, u_m^k, \pi_m, \Pi) \\ & z_{m+1}^0 = \bar{f}_m(z_m^{K_m}, y_m^{K_m}, u_m^{K_m}) \\ & z_0^0 = \hat{z}(\ell) = p \\ & m = 0, \dots, M-1 \quad k = 0, \dots, K_m-1 \end{aligned}$$

27,135 constraints, 9630 LB & UB

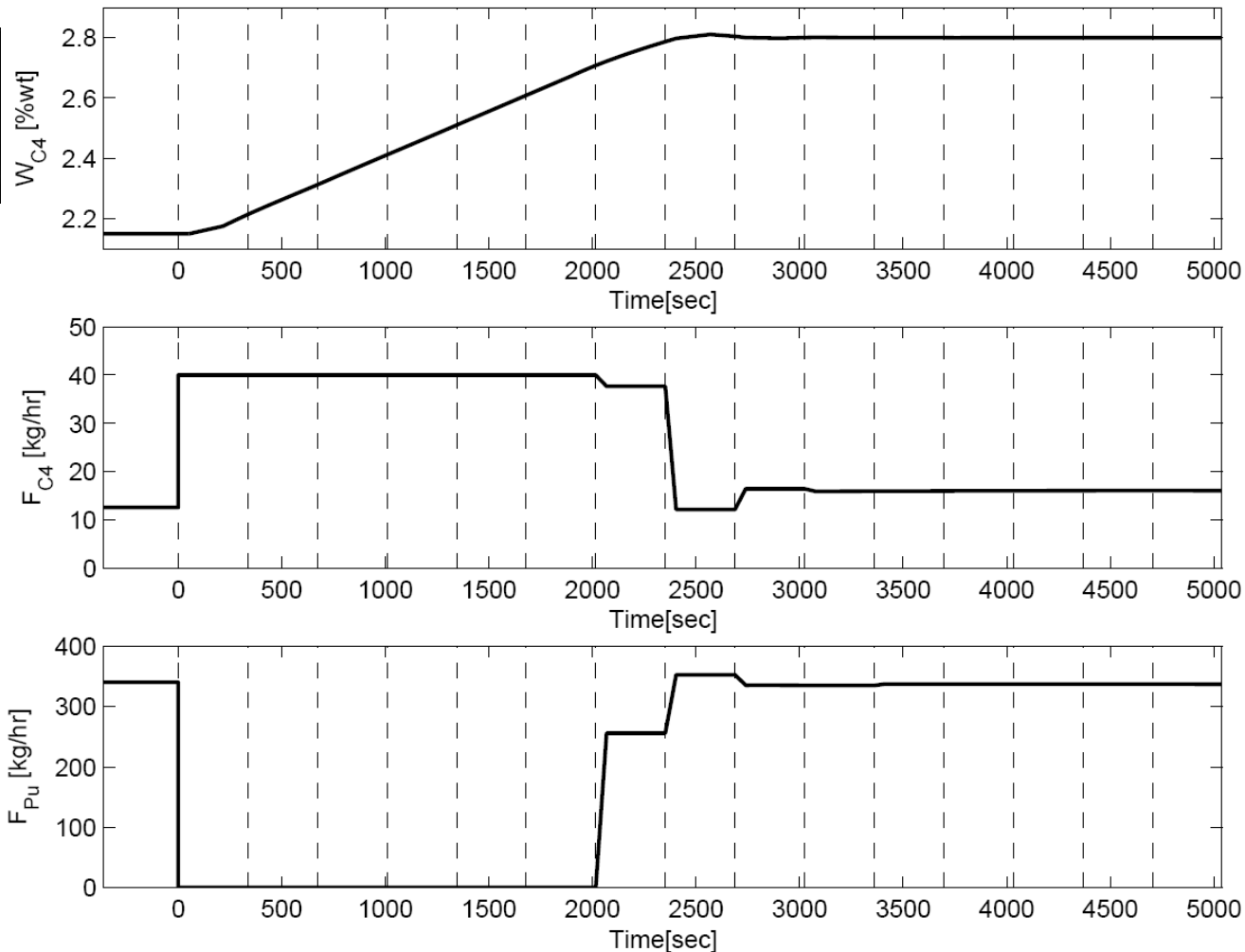


Off-line Solution with IPOPT

Algorithmic Step	CPU(s)
Full Solution (10 iterations)	351.5
Single Factorization of KKT Matrix	33.9
Step Computation (single backsolve)	0.94
Rest of Steps	0.12

**Feedback
Every 6 min**

Nonlinear Model Predictive Control Grade Transition to Reduce MW of LDPE



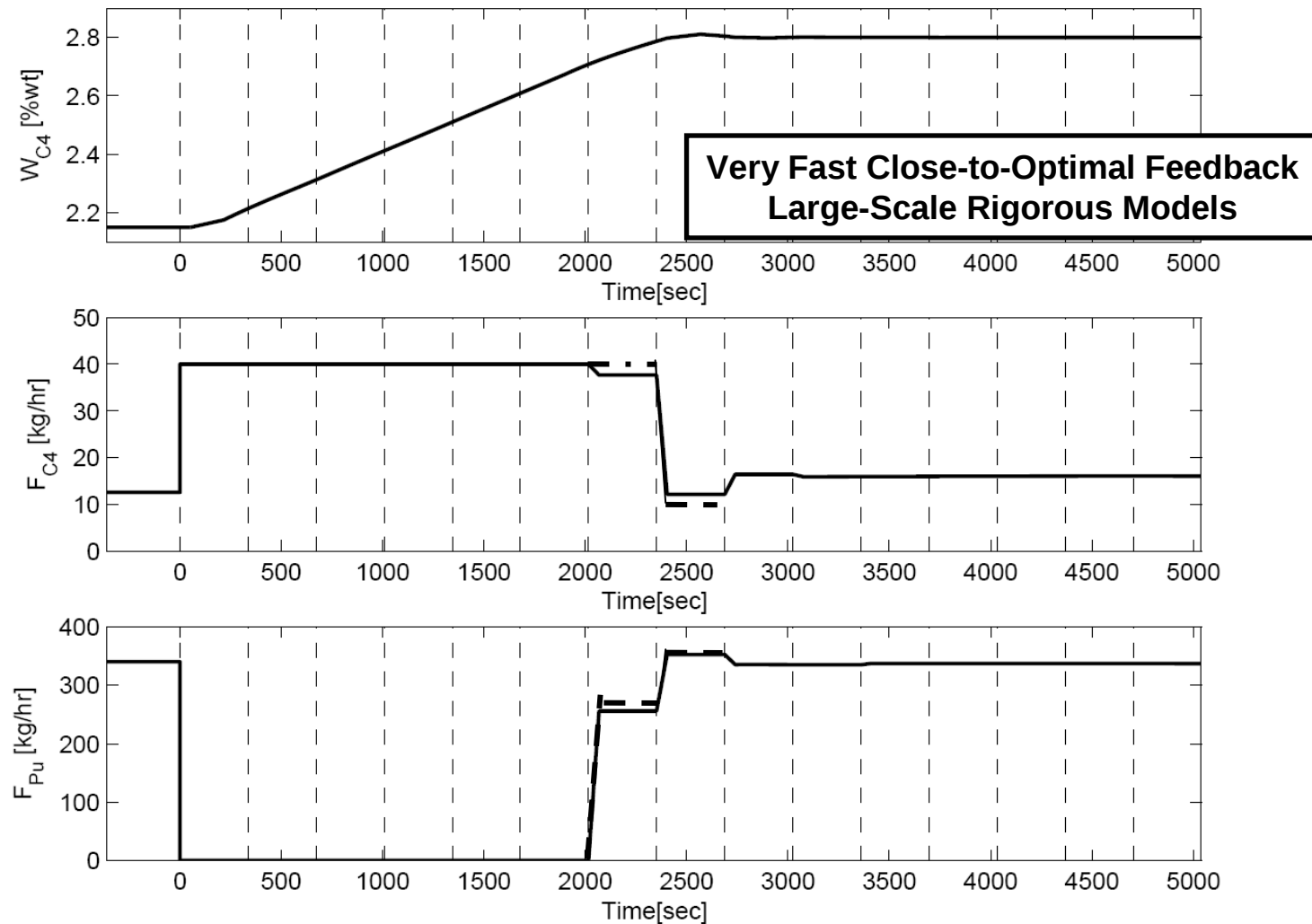
Ideal NMPC controller - computational delay not considered

Time delays as disturbances in NMPC

Optimal Feedback Policy → (On-line Computation 351 CPU s)

Nonlinear Model Predictive Control

- Optimal Policy vs. NLP Sensitivity -Shifted → (On-line Computation 1.04 CPU s)



Analogous Strategy Developed for Moving Horizon Estimation



Summary

Optimization plays a central role in all aspects of chemical process engineering

- Challenging Real-World Applications
- Modeling-design-control-operations
- Conflict between detailed models (off-line analysis) and time critical computation (on-line optimization)

Challenge to develop efficient optimization strategies with multiple model levels

- Closed -- Semi-closed – Fully Open

Advanced algorithms needed to integrate optimization models

- Across process systems
- Across operating time scales
- Across functionalities



Related Papers at this Conference (and more details)

Reduced Order Optimization Models

- 429a “Reduced-Order Model for Dynamic Optimization of Pressure Swing Adsorption” A. Agarwal, L. T. Biegler, S. E. Zitney
- 398c “Advanced Process Engineering Co-Simulation Using CFD-Based Reduced Order Models,” Y-D Lang, L. T. Biegler, S. Munteanu, J. Madsen, S. E. Zitney

Open Optimization Models and Parallel Decomposition

- 149d “Efficient Parallel Solution of Dae Constrained Optimization Problems with Loosely Coupled Algebraic Components,” C. D. Laird, V. M. Zavala, L. T. Biegler
- 531a “Modeling and Optimization of Polymer Electrolyte Membrane Fuel Cells,” P. Jain, L. T. Biegler, M. S. Jhon

Fast NMPC and MHE

- 149b “A Moving Horizon Estimation Algorithm Based on NLP Sensitivity,” V. M. Zavala, L. T. Biegler
- 256a “Stability and Performance Analysis of Nlp Sensitivity-Based NMPC Controllers,” V. M. Zavala, L. T. Biegler



Acknowledgements

Funding

- US Department of Energy
- ExxonMobil Upstream Research Corporation
- US National Science Foundation
- Center for Advanced Process Decision-making (CMU)

Research Colleagues

- Anshul Agarwal
- Parag Jain
- Carl Laird
- Yi-dong Lang
- Andreas Wächter
- Victor Zavala
- Steve Zitney

<http://dynopt.cheme.cmu.edu>