A Moving Horizon Estimation Algorithm Based on NLP Sensitivity

Victor M. Zavala and Lorenz T. Biegler

Department of Chemical Engineering
Carnegie Mellon University

vzavala@cmu.edu

http://dynopt.cheme.cmu.edu

AIChE Annual Meeting
November 6th 2007

Carnegie Mellon

CAPD CENTER

Background and Motivation

Background and Motivation

Full Space NLP Solvers ~100,000 Constraints, ~1000 Degrees of Freedom

Large-Scale NLP

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

KKT Conditions

- Newton's Method + Exact Derivatives

Sparse Factorization of Full KKT Matrix – Most Expensive Step

KKT Matrix

$$\begin{bmatrix} H_k & A_k \\ A_k^T & \mathbf{K^*} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} \nabla_x \mathcal{L}(x_k, \lambda_k) \\ c(x_k) \end{bmatrix}$$

Already Factorized — Newton Step — KKT Conditions

- KKT Matrix Contains Information About Solution
- This Talk: Reuse KKT Matrix

1) Fast Approximate Solutions for Neighboring NLPs
2) Extraction of Reduced Hessian Information
   - Covariance Analysis for Large-Scale Estimation

IPOPT *Waechter, Biegler 2006*

Fast Moving Horizon Estimation

NLP Sensitivity

Fast Approximate Solutions

NLP Sensitivity

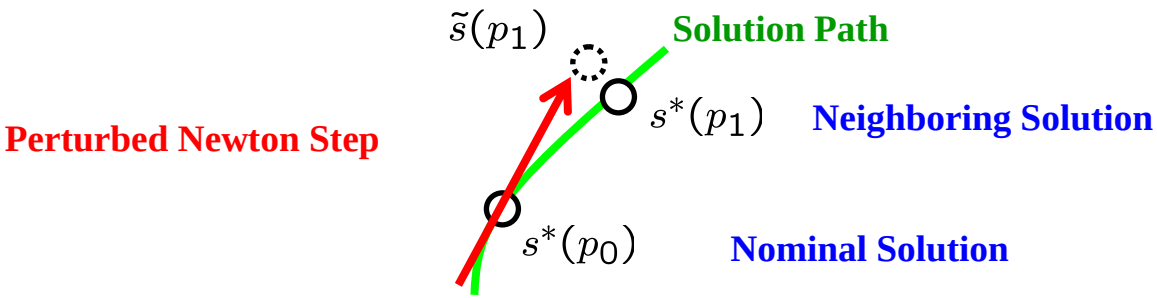
$$\left.\begin{array}{ll} \min & f(x,p) \\ \text{s.t.} & c(x,p) = 0 \\ & x \geq 0 \end{array}\right\} \mathrm{P}(p)$$

Data

Solution Vector

$$s^*(p)^T = [x^{*T}(p)\ \lambda^{*T}(p)]$$

- NLP Sensitivity - Existence and Differentiability of Solution Path - *Fiacco, 1983*



- KKT System at Nominal Solution

Already Factorized

$$\mathbf{K}^* \ \Delta s = -\nabla_s \mathcal{L}(s^*(p_0), p_0)$$

p_1

$(\tilde{s}(p_1) - s^*(p_0))$

- Approximation Error $|\tilde{s}(p_1) - s^*(p_1)| \approx O|p_1 - p_0|^2$

- Fast Backsolve with Factorized KKT Matrix

NLP Sensitivity

- Perturb Optimal Value of Internal Variables *Z., Laird & Biegler, 2006*

$x = [\bar{x}\ y]$ $\quad y^*(p_0) \rightarrow y_d$

Already Factorized

$$\begin{bmatrix} H(p_0) & A(p_0) \\ A(p_0)^T & \mathbf{K^*} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \end{bmatrix} = -\begin{bmatrix} \nabla_x \mathcal{L}(s^*(p_0), p_0) \\ c(x^*(p_0), p_0) \end{bmatrix}$$

Already Factorized

Relax

$$\boxed{\Delta\sigma} + \cancel{\left[\begin{array}{cc|c} H_{\bar{x}\bar{x}} & H_{\bar{x}y} & A_{\bar{x}} \\ H_{y\bar{x}} & \mathbf{K^*} & A_y \\ \hline A_{\bar{x}}^T & A_y^T & \end{array}\right] \left[\begin{array}{c} \Delta\bar{x} \\ \Delta y \\ \hline \Delta\lambda \end{array}\right] = -\left[\begin{array}{c} \nabla_{\bar{x}} \mathcal{L}(s^*(p_0), p_0) \\ \nabla_y \mathcal{L}(s^*(p_0), p_0) \\ \hline c(\bar{x}^*(p_0), y^*(p_0)) \end{array}\right]}$$

$$\boxed{\Delta y \ = \ -(y_d - y^*(p_0))}$$

Fix

- Augmented KKT System

$$\left[\begin{array}{c|c} \mathbf{K^*} & \mathbf{E}_\sigma \\ \hline \mathbf{E}_y^T & 0 \end{array}\right] \left[\begin{array}{c} \Delta s \\ \hline \Delta\sigma \end{array}\right] = -\left[\begin{array}{c} \nabla_s \mathcal{L}(s^*(p_0), p_0) \\ \hline y_d - y^*(p_0) \end{array}\right]$$

- Approximation Error $|\tilde{s}(y_d) - s^*(y_d)| \approx O|y_d - y^*(p_0)|^2$

- Reuse Factorization with Schur Complement Technique

Reduced Hessian Information

Analysis of Second Order Conditions

Reduced Hessian

-KKT System at Solution

Full Hessian

$$x = [\bar{x}\ u]$$

$$\begin{bmatrix} H & A \\ A^T & \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = - \begin{bmatrix} \nabla_x \mathcal{L}(s^*) \\ c(x^*) \end{bmatrix}$$

- Split Variables into Basic and Non-Basic (Degrees of Freedom)

$$\left[\begin{array}{cc|c} H_{\bar{x}\bar{x}} & H_{\bar{x}u} & A_{\bar{x}} \\ H_{u\bar{x}} & H_{uu} & A_u \\ \hline A_{\bar{x}}^T & A_u^T & \end{array}\right] \left[\begin{array}{c} \bar{X} \\ U \\ \hline \Lambda \end{array}\right] = - \left[\begin{array}{c} 0 \\ \mathbf{I}_{n_u} \\ \hline 0 \end{array}\right]$$

$\mathbf{K}^*$

- Reduced Hessian $Z^T H Z$ **Full Hessian Projected onto the Null Space Matrix** $Z = \begin{bmatrix} -A_{\bar{x}}^{-1} A_u \\ \mathbf{I}_u \end{bmatrix}$

Never Formed in IPOPT

- $U = (Z^T H Z)^{-1}$ *Z. & Biegler, 2007*

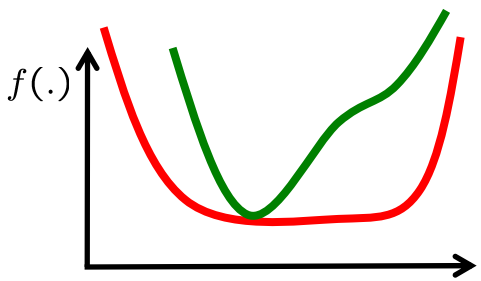
- Reduced Hessian Can be Extracted Through Fast Backsolves

- Large-Scale Applications:

- Analysis Second Order Conditions

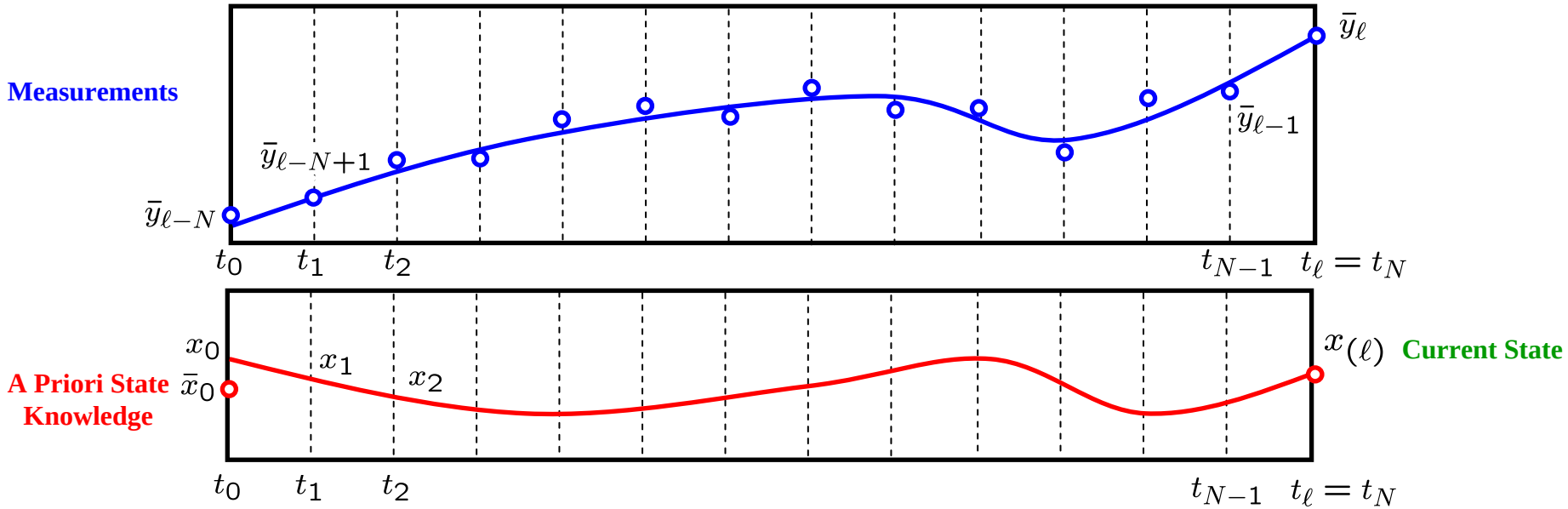
- Inference Parameter Estimation $\mathbf{V}_\theta \approx (Z^T H Z)^{-1}$ *Bard, 1974*

Covariance Parameters



FAST Moving Horizon Estimation

Moving Horizon Estimation



A Priori Covariance

$$\min_{\boxed{x_0, w_k}} (x_0 - \bar{x}_0)^T \mathbf{\Pi}_0^{-1} (x_0 - \bar{x}_0) + \sum_{k=0}^{N} L_k(w_k, v_k)$$

Deviation Costs

$$s.t. \quad \begin{aligned} x_{k+1} &= f_k(x_k, w_k) && \text{Dynamic Model} \\ \text{Measurements} \longrightarrow \bar{y}_{k+\ell-N} &= h_k(x_k) + v_k && \text{Output Model} \end{aligned}$$

$\mathbf{P}(\ell)$

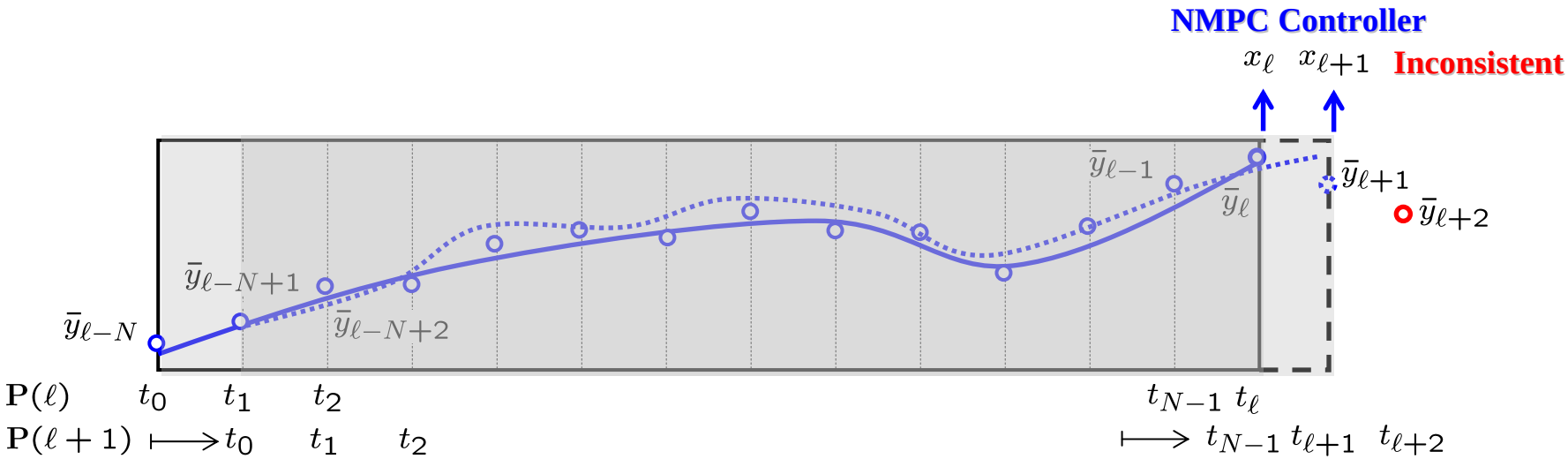
Large State Dimensionality $100 - 1,000$

Many Degrees of Freedom $(100 - 1,000) + (10 - 100)N$

Highly Nonlinear, Ill-Posed

Solution Time ~ Several Minutes

Fast Moving Horizon Estimation



$$\underbrace{\begin{aligned} \min_{x_0, w_k} \quad & (x_0 - \bar{x}_0^\ell)^T \mathbf{\Pi}_0^{-1} (x_0 - \bar{x}_0^\ell) + \sum_{k=0}^{N} L_k(v_k, w_k) \\ s.t. \quad & \\ x_{k+1} &= f_k(x_k, w_k) \\ \boxed{\bar{y}_{k+\ell-N}} &= h_k(x_k) + v_k \end{aligned}}_{\mathbf{P}(\ell)}$$

Computational Delay **– Stability NMPC**

NLPs are *Parametric*

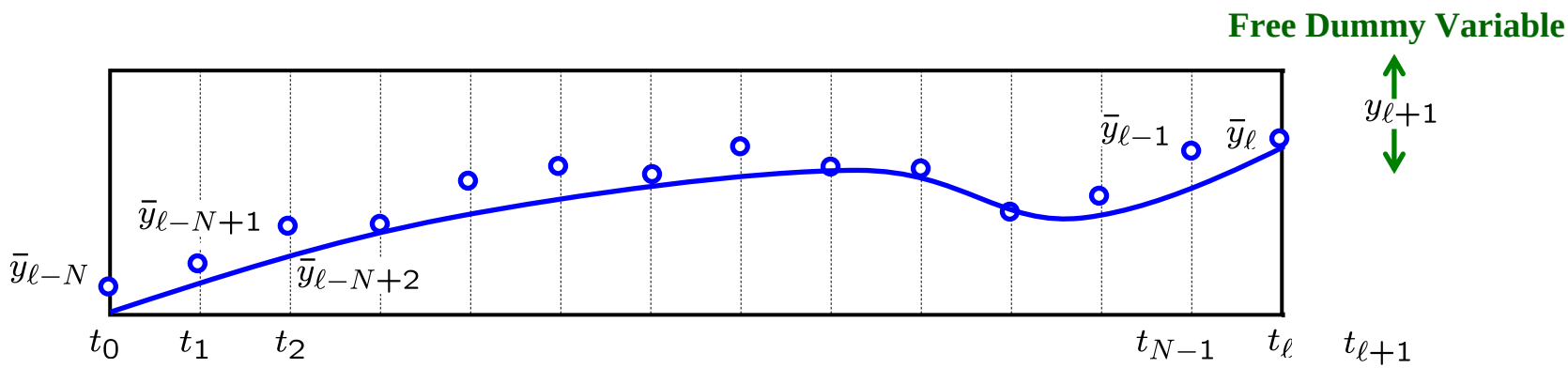
Solve Reference Problem Between Sampling Times

Fast On-Line Approximation with Sensitivity

$$\mathbf{P}(\ell+1) \left\{ \begin{aligned} \min_{x_0, w_k} \quad & (x_0 - \bar{x}_0^{\ell+1})^T \mathbf{\Pi}_0^{-1} (x_0 - \bar{x}_0^{\ell+1}) + \sum_{k=0}^{N} L_k(v_k, w_k) \\ s.t. \quad & \\ x_{k+1} &= f_k(x_k, w_k) \\ \boxed{\bar{y}_{k+\ell-N+1}} &= h_k(x_k) + v_k \end{aligned} \right.$$

Fast Moving Horizon Estimation

- Sensitivity-Based Shifting Strategy *Z.; Laird & Biegler, 2007*



1) Solve *Extended* Problem $\bar{\mathbf{P}}(\ell)$

$$\min_{x_0, w_k, \boxed{y_{\ell+1}}} \; (x_0 - \bar{x}_0^\ell)^T \mathbf{\Pi}_0^{-1} (x_0 - \bar{x}_0^\ell) + \sum_{k=0}^{N} L_k(v_k, w_k) \boxed{+ L_{N+1}(v_{N+1}, w_{N+1})}$$

$$\bar{\mathbf{P}}(\ell) \quad s.t. \quad \begin{aligned} x_{k+1} &= f_k(x_k, w_k), & k &= 0, \ldots, N-1 \\ \bar{y}_{k+\ell-N} &= h_k(x_k) + v_k, & k &= 0, \ldots, N \\ \boxed{\begin{aligned} x_{N+1} &= f_N(x_N, w_N) \\ y_{\ell+1} &= h_{N+1}(x_{N+1}) + v_{N+1} \end{aligned}} \end{aligned}$$

Dummy Variable at Solution = Output Model Extrapolation

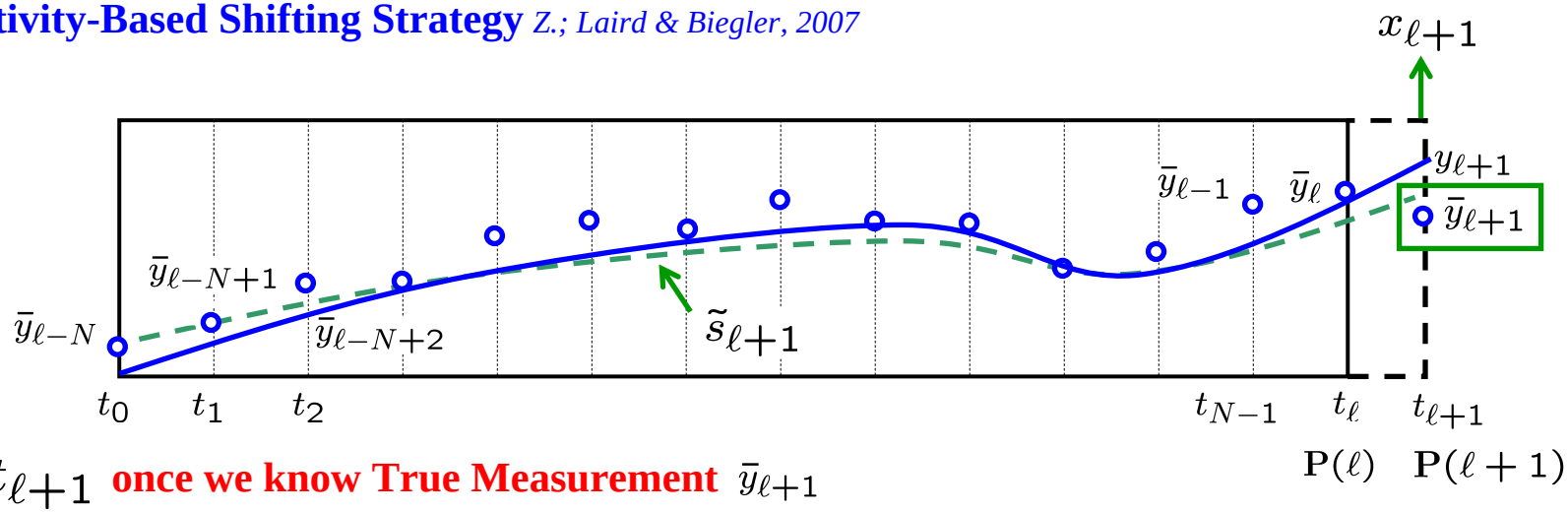
$$y^*_{\ell+1} = h(x_N)$$

Hold KKT System of $\bar{\mathbf{P}}(\ell)$ **at** s^*_ℓ

$$\mathbf{K} \Delta s = 0$$

Fast Moving Horizon Estimation

- Sensitivity-Based Shifting Strategy *Z.; Laird & Biegler, 2007*



2) At $t_{\ell+1}$ once we know True Measurement $\bar{y}_{\ell+1}$

Perturb Dummy *Variable* to True Measurement $y^*_{\ell+1} \rightarrow \bar{y}_{\ell+1}$

NLP Sensitivity Fix-Relax Strategy

$$\left[\begin{array}{c|c} \mathbf{K} & \mathbf{E}_\sigma \\ \hline \mathbf{E}_y^T & 0 \end{array}\right] \left[\begin{array}{c} \Delta s \\ \hline \Delta\sigma \end{array}\right] = -\left[\begin{array}{c} 0 \\ \hline \bar{y}_{\ell+1} - y^*_{\ell+1} \end{array}\right]$$

Fast On-line Computation

$$\tilde{s}_{\ell+1} = s^*_\ell + \Delta s$$

Accurate Approximation

Fast Moving Horizon Estimation

Extract Covariance of State Estimate Π_N Never Formed in IPOPT

$$\min_{x_0, w_k} \quad (x_0 - \bar{x}_0^\ell)^T \Pi_0^{-1} (x_0 - \bar{x}_0^\ell) + \sum_{k=0}^{N} L_k(v_k, w_k)$$

$$s.t.$$

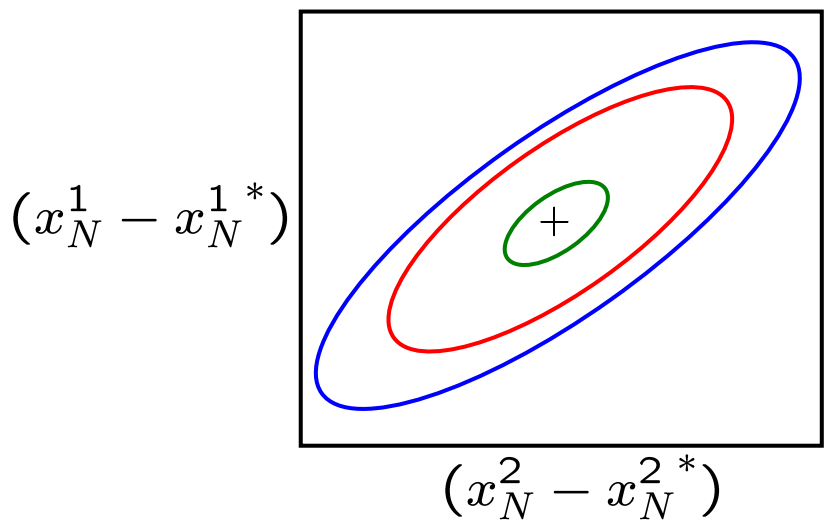
$$\begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ \bar{y}_{k+\ell-N} &= h_k(x_k) + v_k \end{aligned}$$

Large-Scale MHE

$$\longleftrightarrow$$

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

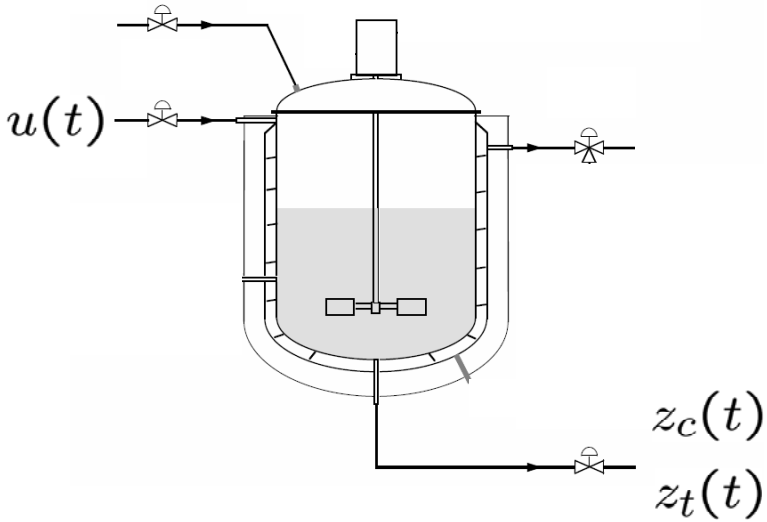
IPOPT



- Π_N **Normally Computed through Kalman Filter Recursion**

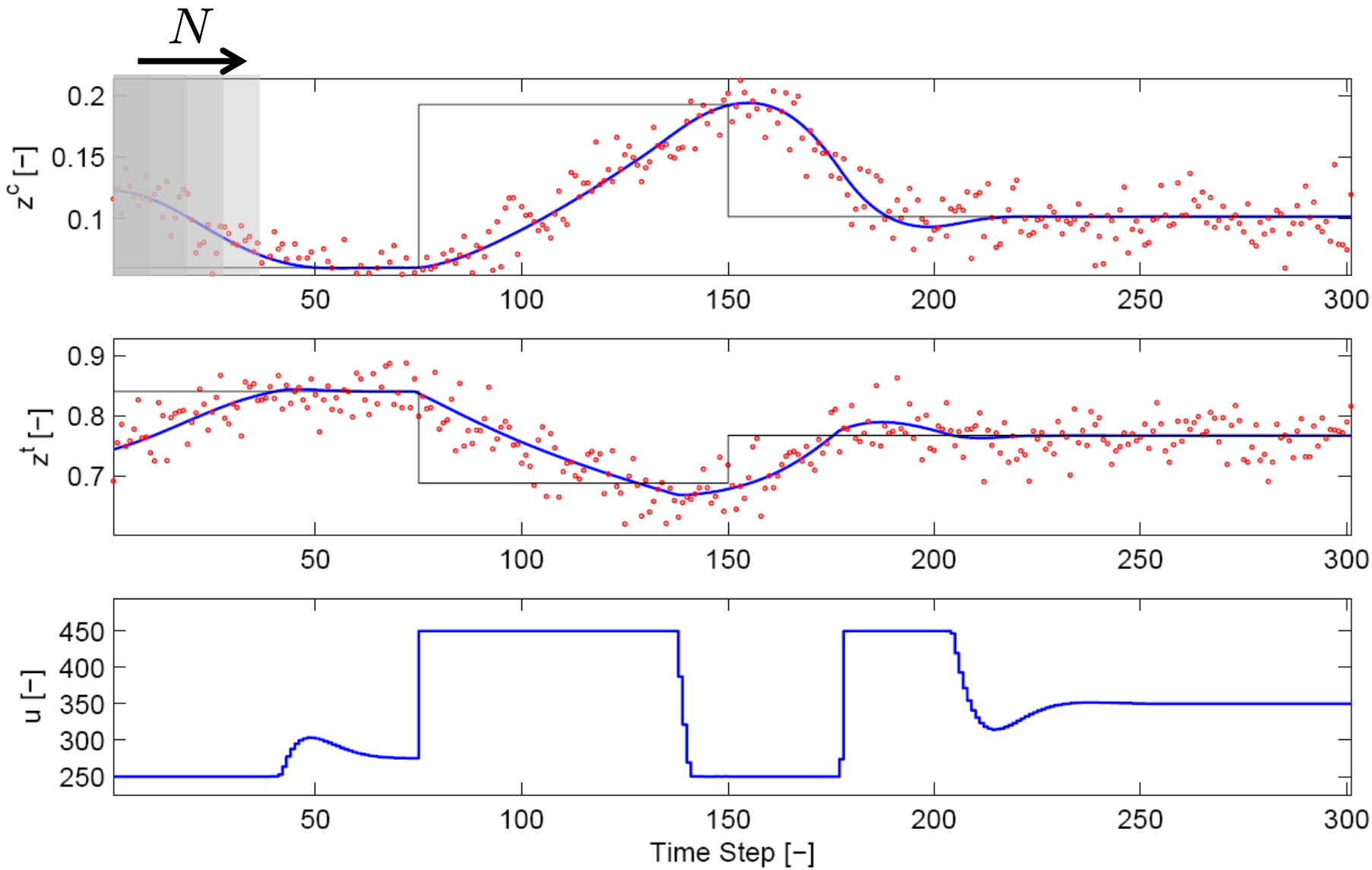
- Π_N **Is Not Reduced Hessian but Can be Extracted with Reduced Hessian Tool**

Case Study - CSTR



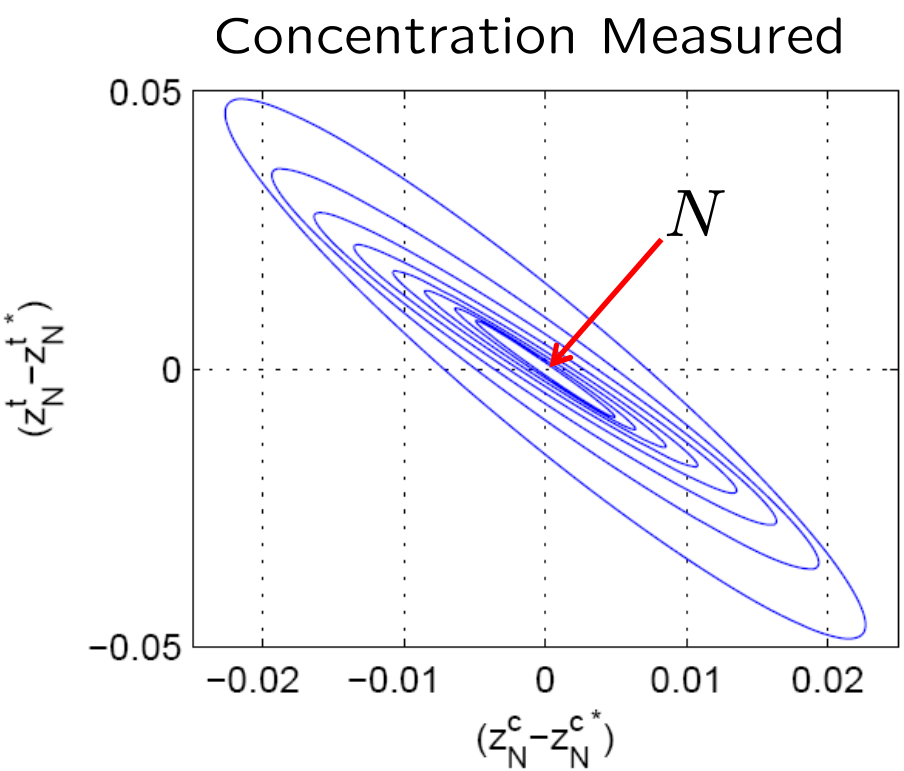
CSTR

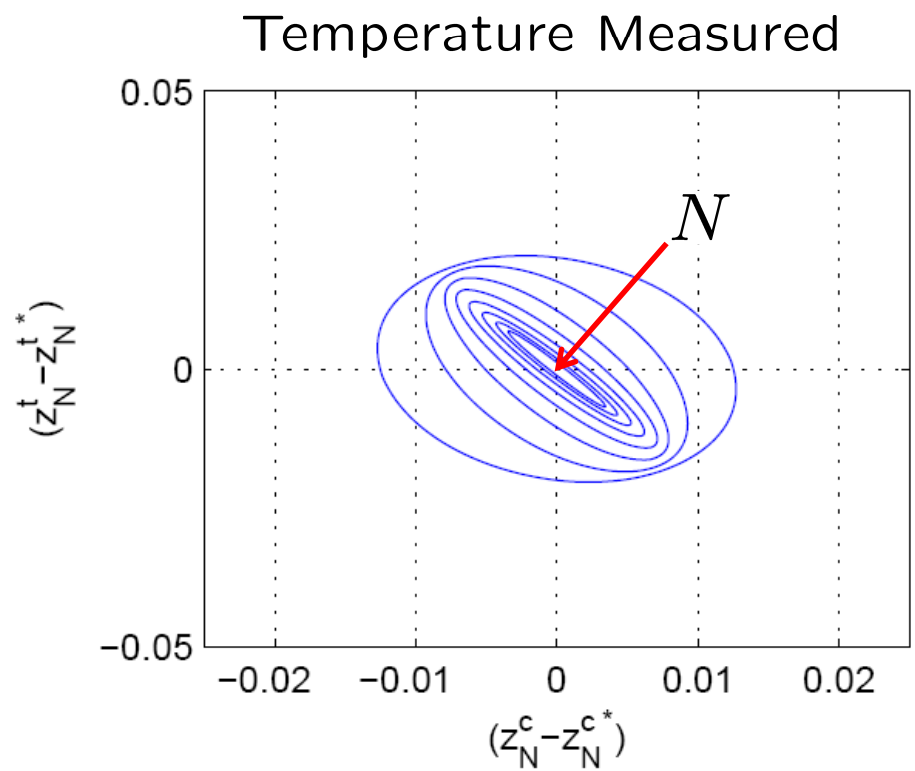
Measurement Data Generated from NMPC Scenario



CSTR

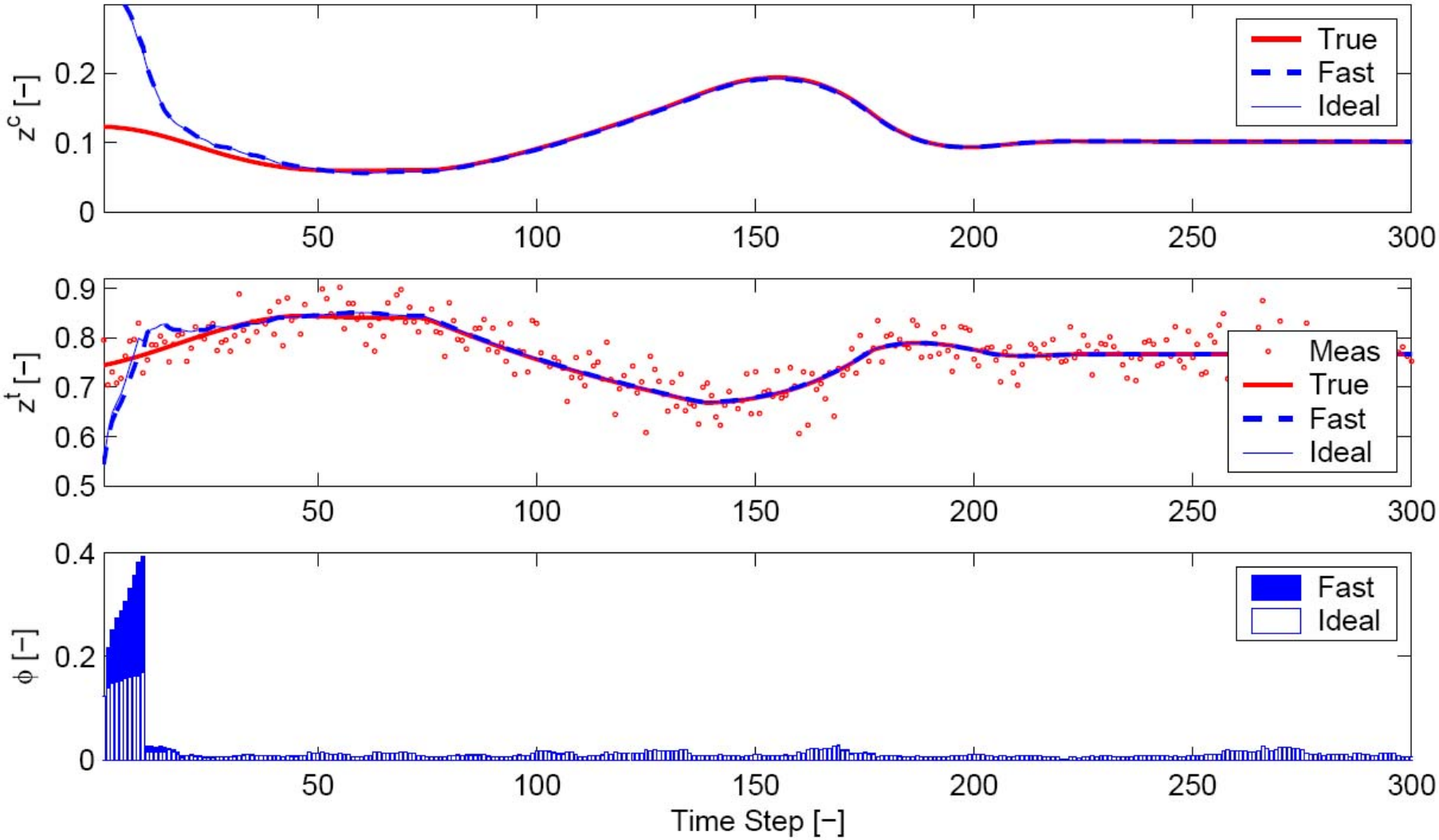
95% Confidence Regions for $N = 5, 10, 15, \ldots, 30$





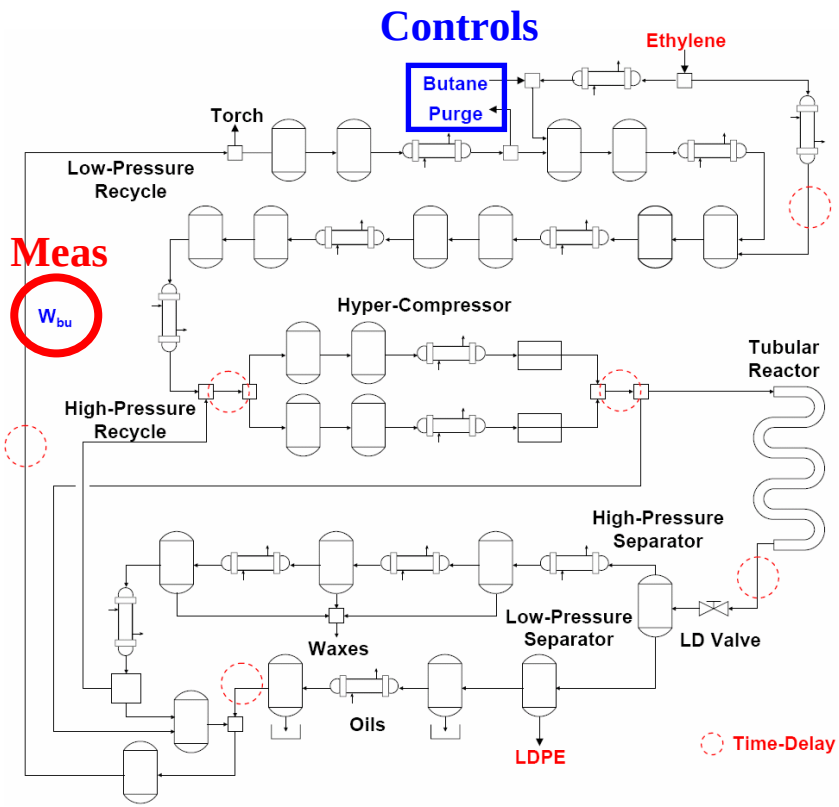
CSTR

Performance of Fast MHE – Effect of NLP Sensitivity Errors



Case Study – Industrial Low-Density Polyethylene Process

LDPE Process



Dynamic Model ~ 350 DAEs
Sampling Time = 6 min
Estimation Horizon N = 15

NLP IPOPT

$$\min_{x_0, w_k} \quad (x_0 - \bar{x}_0^\ell)^T \Pi_0^{-1} (x_0 - \bar{x}_0^\ell) + \sum_{k=0}^{N} L_k(v_k, w_k)$$

$$s.t. \quad \begin{aligned} x_{k+1} &= f_k(x_k, w_k) \\ \bar{y}_{k+\ell-N} &= h_k(x_k) + v_k \end{aligned}$$

27,121 Constraints, 9330 Bounds
294 Degrees of Freedom +1 Dummy

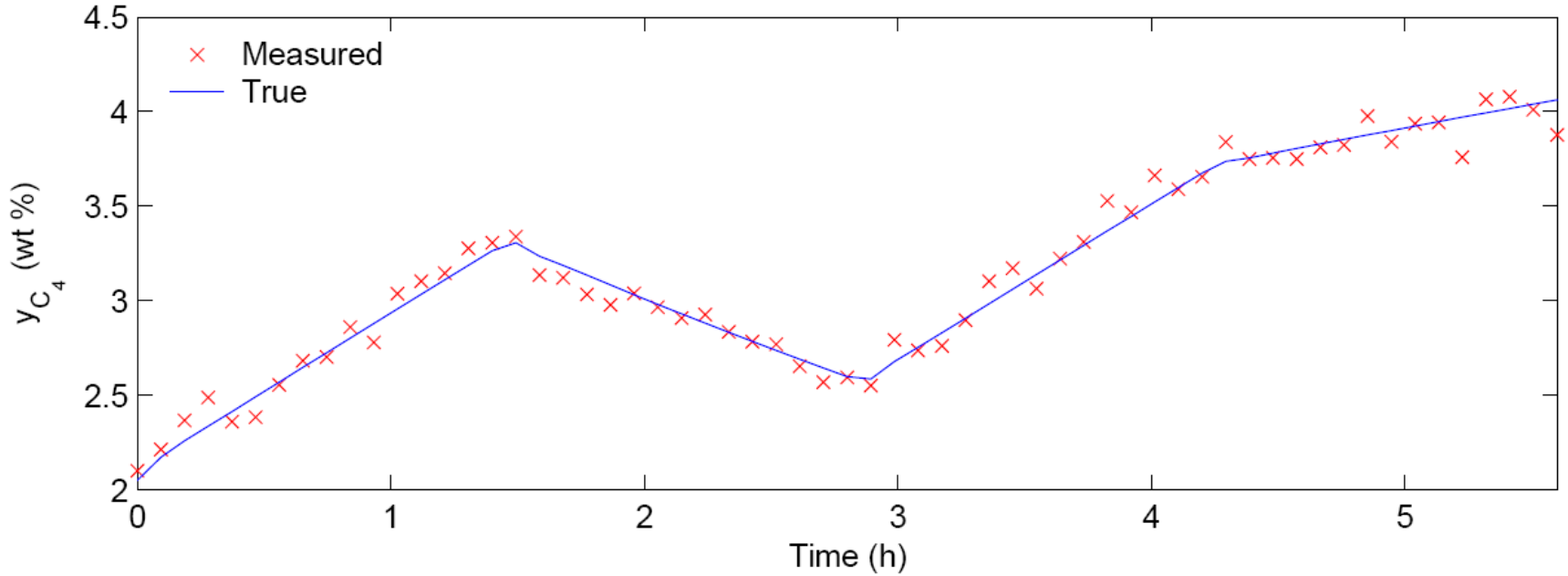
Algorithmic Step	CPUs
Full Solution (6 iterations)	202.64
Single Factorization of KKT Matrix	33.77
Step Computation (single backsolve)	0.9-1.0
Rest of Steps	0.936

← **On-Line Cost Standard MHE**

← **On-Line Cost Fast MHE**

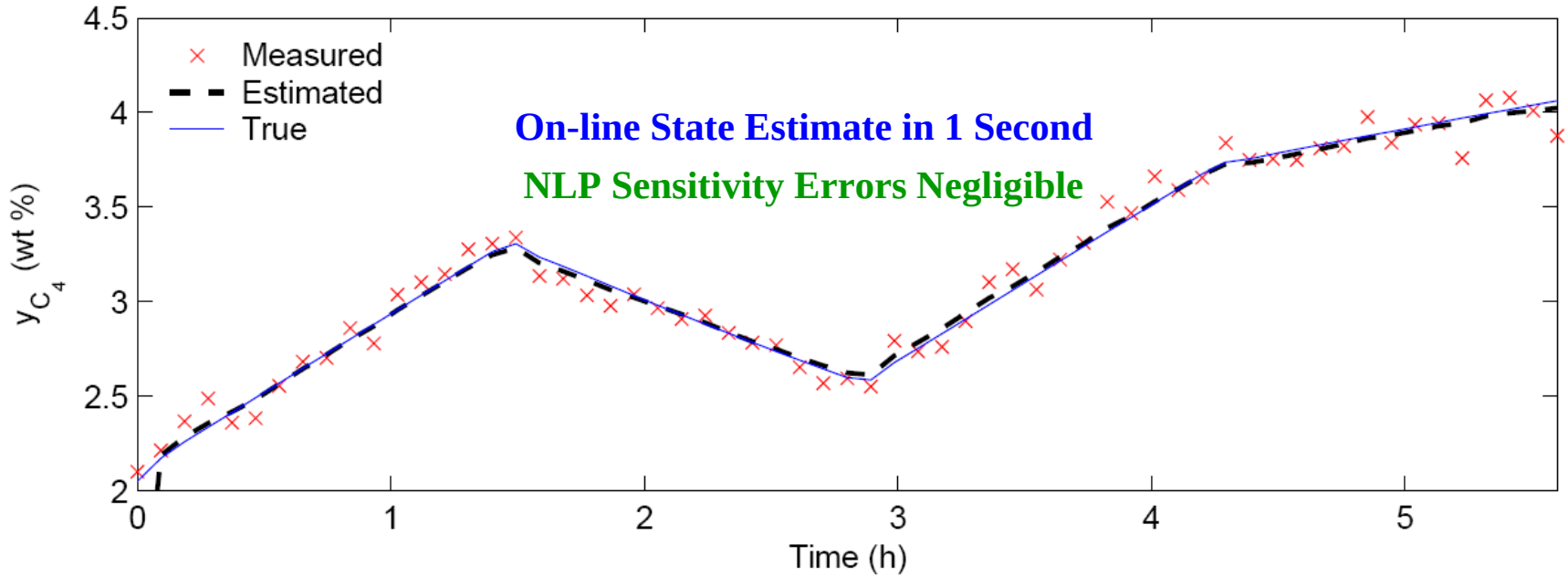
LDPE Process

Measurement Profiles Gaussian Noise 5% SD



LDPE Process

Measurement Profiles Gaussian Noise 5% SD



Conclusions and Future Work

Conclusions and Future Work

- Reuse KKT Matrix at Solution
  - NLP Sensitivity Calculations
  - Reduced Hessian Information
  - IPOPT

- Fast Moving Horizon Estimation
  - Industrial-Size MHE without On-Line Computational Delay

-Future Work
  - Fix-Relax to Convert KKT Matrix into Fast QP Active-Set Solver
  - Warm-Start for NLP Sequences: MINLP

A Moving Horizon Estimation Algorithm Based on NLP Sensitivity

Victor M. Zavala and Lorenz T. Biegler

Department of Chemical Engineering
Carnegie Mellon University

vzavala@cmu.edu

http://dynopt.cheme.cmu.edu

AIChE Annual Meeting
November 6th 2007

Carnegie Mellon

CAPD CENTER