A Fast Computational Framework for Large-Scale Moving Horizon Estimation

Victor M. Zavala, Carl D. Laird and Lorenz T. Biegler

Department of Chemical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

vzavala@cmu.edu

http://dynopt.cheme.cmu.edu

8th International Symposium on Dynamics and Control of Process Systems
June 8th 2007

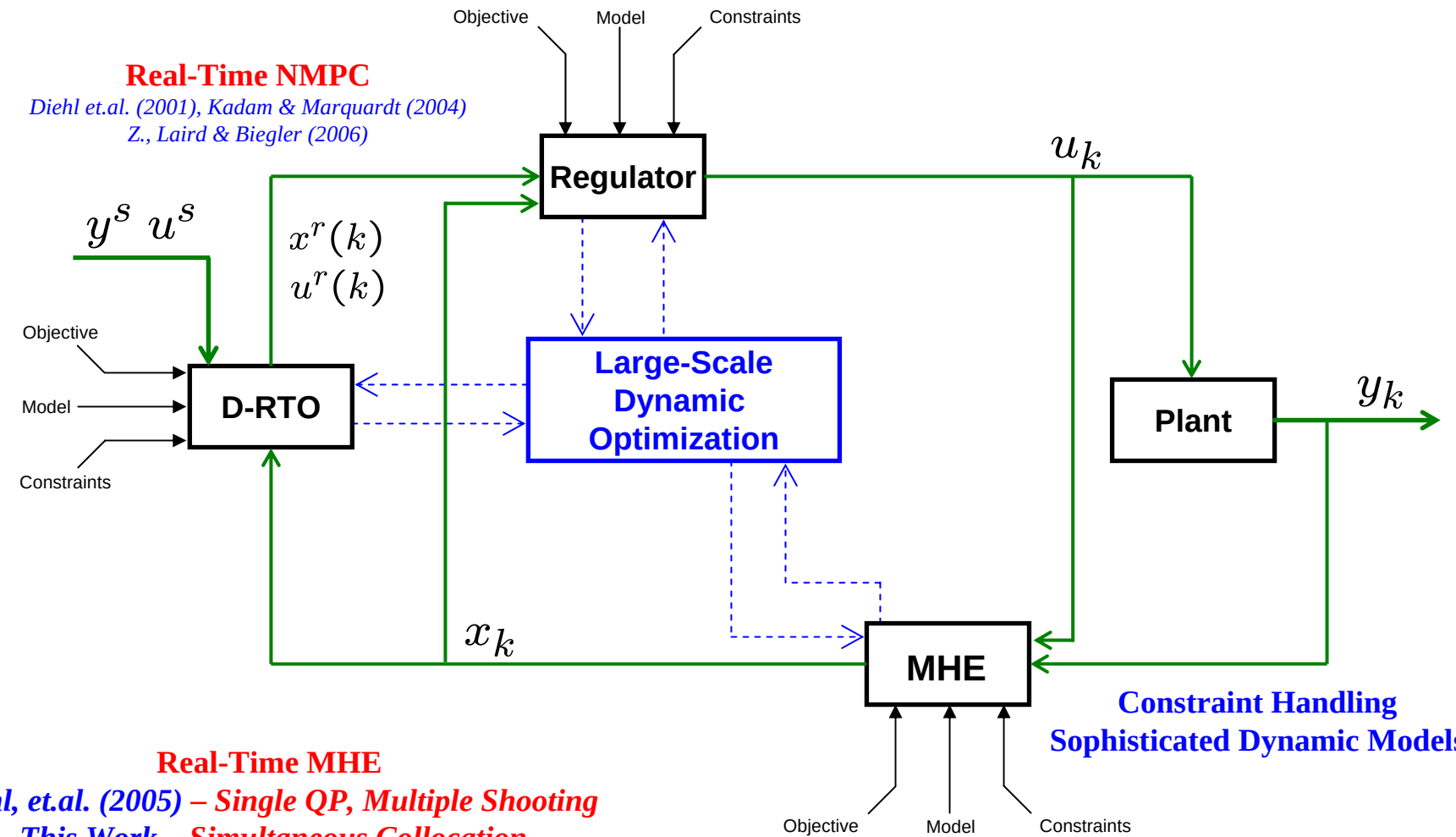
Carnegie Mellon

CAPD CENTER

Motivation

Complex Process Operations – Systematic Decision-Making

Moving Horizon Estimation



Real-Time NMPC

Diehl et.al. (2001), Kadam & Marquardt (2004)
Z., Laird & Biegler (2006)

Constraint Handling
Sophisticated Dynamic Models

Real-Time MHE

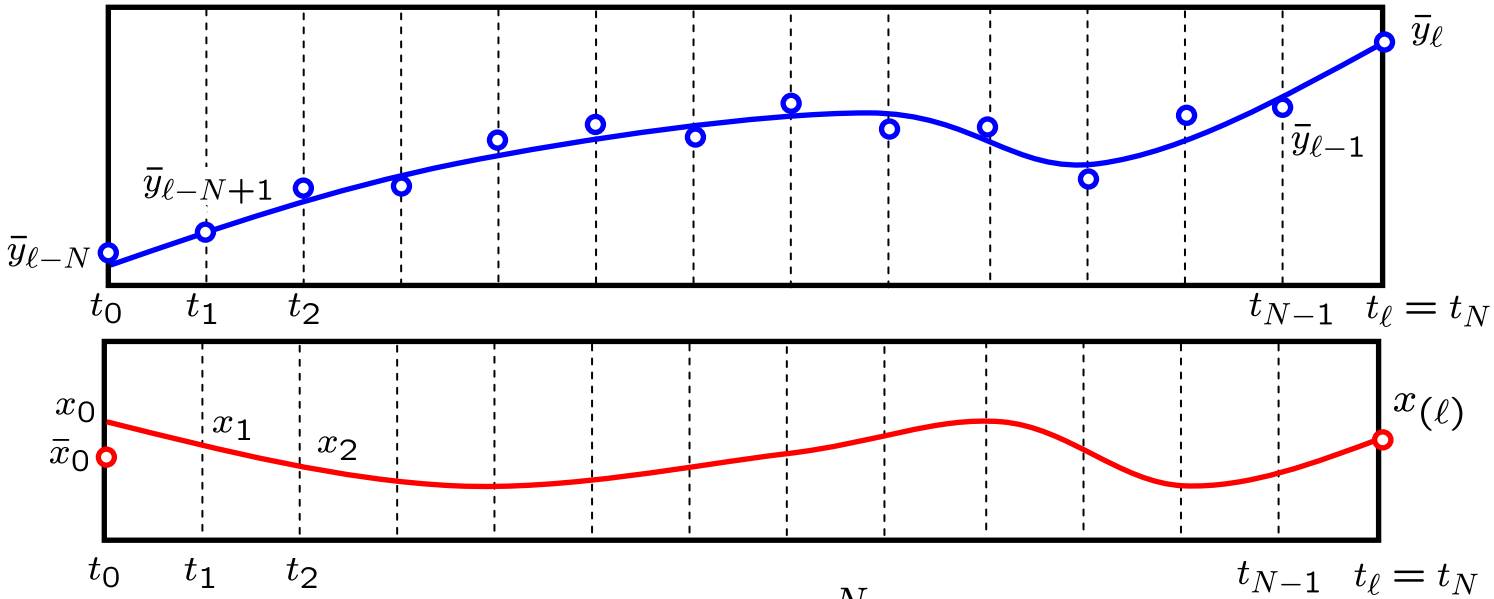
Diehl, et.al. (2005) – Single QP, Multiple Shooting
This Work – Simultaneous Collocation

MHE – Theoretical Developments

Robertson, Lee and Rawlings (1994), Albuquerque and Biegler (1996),
Rao, Rawlings and Mayne (2003), Rawlings and Bakshi (2006)

Moving Horizon Estimation Problem

Moving Horizon Estimation



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

$$s.t. \quad \begin{array}{rcll} x_{k+1} & = & f_k(x_k, w_k) & k = 0, ..., N-1 \\ \bar{y}_{k+\ell-N} & = & h_k(x_k) + v_k & k = 0, ..., N \\ x_k & \in & \mathcal{X}_k, \; w_k \in \mathcal{W}_k, \; v_k \in \mathcal{V}_k & \end{array} \qquad \mathbf{P}(\ell)$$

Large State Dimensionality $dim(x_k) \approx 100 - 10,000$

Degrees of Freedom $dim(x_0) + dim(w_k)N \approx (100 - 10,000) + (10 - 100)N$

Highly Nonlinear, Ill-Posed

Solution Time → *Order of Minutes*

Solved as Continuous-Time DAE-Constrained Problem

Moving Horizon Estimation

$$\min_{x_0,\, w(t)} \quad \int_{t_0}^{t_f} \varphi\left(x_d(t), x_a(t), w(t)\right) dt$$

s.t.

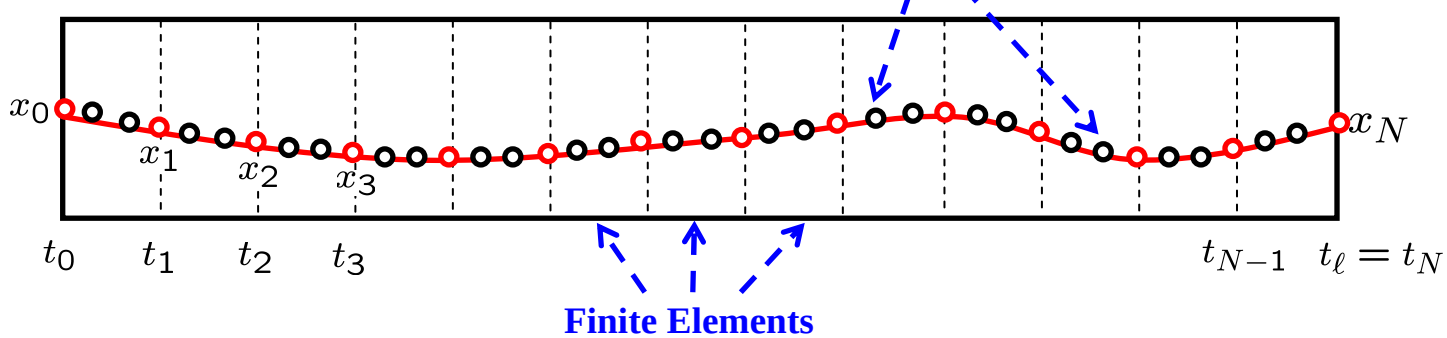
$$\left.\begin{aligned} \frac{dx_d}{dt} &= f\left(x_d(t), x_a(t), w(t)\right) \\ 0 &= g\left(x_d(t), x_a(t), w(t)\right) \end{aligned}\right\}$$

$$0 \leq h\left(x_d(t), x_a(t), w(t)\right)$$

Simultaneous Collocation Approach

Algebraic Representation of Continuous-Time Dynamic Model

Polynomials
At Collocation Points



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

Large-Scale and Sparse NLP

1) Sparse Linear Algebra in NLP Solver

2) Inexpensive Exact First and Second Order Derivatives

Relies on Efficient NLP Solvers

Moving Horizon Estimation

Full Space NLP Solvers – Knitro, LOQO, *IPOPT* *Wächter, Biegler 2006*

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned} \quad \text{NLP} \quad \longrightarrow \quad \begin{aligned} \min \quad & \varphi_{\mu_\ell}(x) = f(x) - \mu_\ell \sum_{i=1}^{nx} \ln\left(x^{(i)}\right) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned} \quad \text{BP}$$

Solve Sequence of BPs with $\mu_\ell \to 0$

$$\begin{aligned} \nabla_x f(x) + \nabla_x c(x)\,\lambda - \nu &= 0 \\ c(x) &= 0 \\ XVe - \mu_\ell\, e &= 0 \end{aligned} \xrightarrow[\text{Exact Derivatives}]{\text{Newton's Method}}$$

Sparse Factorization of KKT Matrix

$$\begin{bmatrix} H_k & A_k & -I \\ A_k^T & 0 & 0 \\ V_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = -\begin{bmatrix} \nabla f(x_k) + A_k\lambda_k - \nu_k \\ c(x_k) \\ X_k V_k e - \mu_\ell e \end{bmatrix}$$

Complexity of Sparse Factorization

$$O(N(n_x + n_w))^\beta \quad \beta = [1, 2]$$

Complexity of Riccati Decomposition

$$O(N(n_x + n_w)^3)$$

Simultaneous Collocation Approach + Full Space NLP Solver

Favorable Computational Complexity over Multi-Shooting and Sequential *-Z. Laird & Biegler 2006*

Moving Horizon Estimation

Full Space NLP Solvers – Knitro, LOQO, *IPOPT* *Wächter, Biegler 2006*

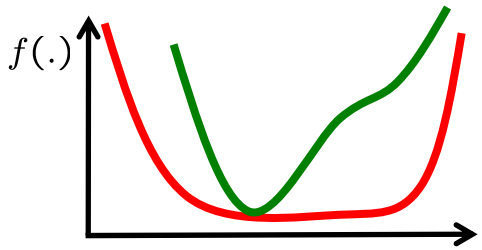
$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x) = 0 \\ & x \geq 0 \end{aligned} \quad \text{NLP} \quad \longrightarrow \quad \begin{aligned} \min \quad & \varphi_{\mu_\ell}(x) = f(x) - \mu_\ell \sum_{i=1}^{nx} \ln\left(x^{(i)}\right) \\ \text{s.t.} \quad & c(x) = 0 \end{aligned} \quad \text{BP}$$

Solve Sequence of BPs with $\mu_\ell \to 0$

$$\begin{aligned} \nabla_x f(x) + \nabla_x c(x)\,\lambda - \nu &= 0 \\ c(x) &= 0 \\ XVe - \mu_\ell e &= 0 \end{aligned} \quad \xrightarrow[\text{Exact Derivatives}]{\text{Newton's Method}}$$

Sparse Factorization of KKT Matrix

$$\begin{bmatrix} H_k & A_k & -I \\ A_k^T & 0 & 0 \\ V_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \nu \end{bmatrix} = -\begin{bmatrix} \nabla f(x_k) + A_k \lambda_k - \nu_k \\ c(x_k) \\ X_k V_k e - \mu_\ell e \end{bmatrix}$$



-Benefits of Exact Derivatives **GAMS, AMPL, ADOL-C, Adifor**

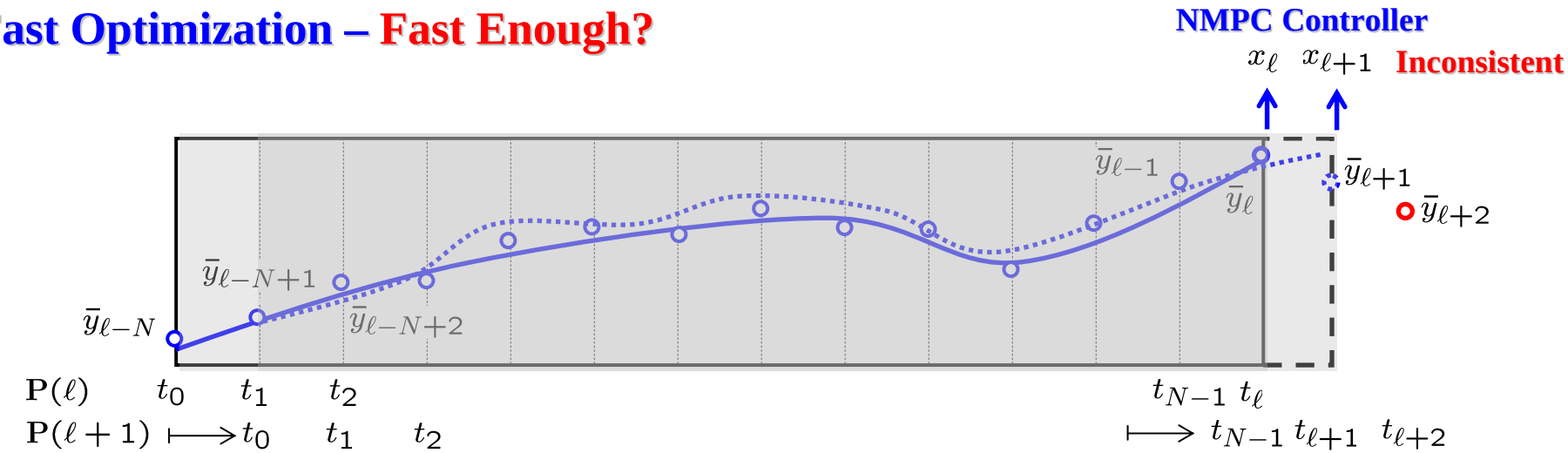
- **Fast Local Convergence**
- **Convergence Independent of Degrees of Freedom vs. Quasi-Newton**
- **Verify Second Order Conditions ~ Observability, Uniqueness** *Z. & Biegler, 2006*
- **Exact Covariance Information ~ Analysis Complex Nonlinear Models**

Simultaneous Collocation Approach + Full Space NLP Solver = Fast Solution of MHE

FAST Moving Horizon Estimation

Fast Moving Horizon Estimation

Fast Optimization – Fast Enough?



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

Computational Delay **– Affects NMPC**

NLPs are *Parametric*

Solve $\mathbf{P}(\ell)$ Between Sampling Times

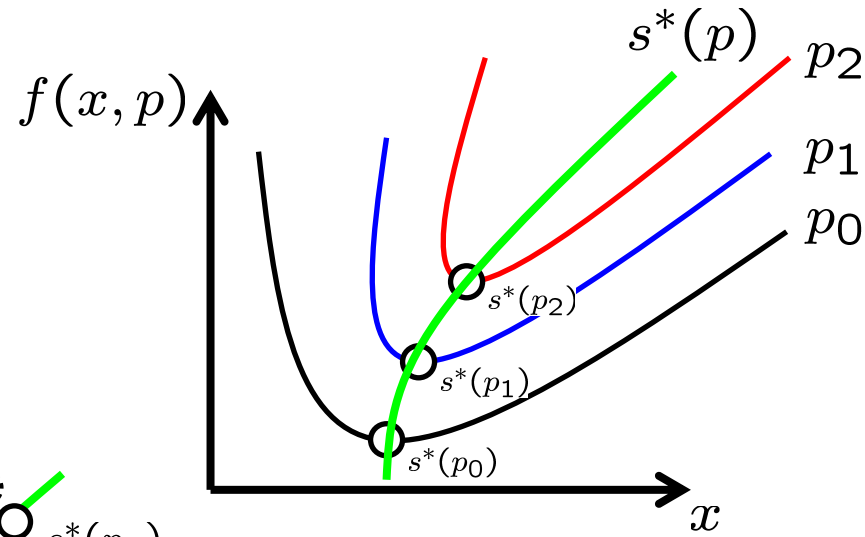
Perturb Solution for Fast *Approximation* $\mathbf{P}(\ell+1)$

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

Large-Scale NLP Sensitivity

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

↑ **MHE Data**

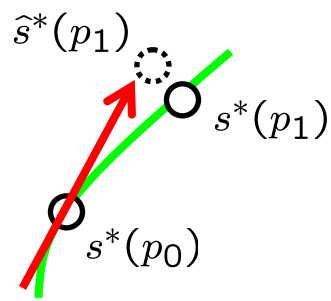


Solution Triplet

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

Optimality Conditions of $\mathrm{P}(p)$

$$\begin{aligned} \nabla_x f(x,p) + \nabla_x c(x,p)\,\lambda - \nu &= 0 \\ c(x,p) &= 0 \\ XVe &= 0 \end{aligned}$$



$$\hat{s}^*(p_1) \approx s^*(p_0) + \left.\frac{\partial s}{\partial p}^T\right|_{p_0} (p - p_0)$$

NLP Sensitivity - Existence and Differentiability of $s^*(p)$ **and** $\frac{\partial s}{\partial p}$ **Unique** - *Fiacco, 1983*

- Main Idea: - Approximation by Taylor Series Expansion - $O(||p - p_0||^2)$

- Equivalent to Newton Step

- Issues: - ***Second Order Conditions*** **hold at** $s^*(p_0)$ **- Observability?**

- Continuity Implies ***Active Set*** **at** $s^*(p_0)$ **Valid in a Finite Neighborhood**

- Keep $||p - p_0||$ **Small**

Fast Moving Horizon Estimation

How to Keep $||p - p_0||$ **Small ?**

- Direct Approach

1) Solve Between t_ℓ **and** $t_{\ell+1}$

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

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

$\mathbf{P}(\ell)$

$$\begin{aligned} \min \quad & f(x, p_0) \\ \text{s.t.} \quad & c(x, p_0) = 0 \\ & x \geq 0 \end{aligned}$$

$\mathbf{P}(p_0)$

Obtain $s^*(p_0)$ $\dfrac{\partial s}{\partial p}$

2) At $t_{\ell+1}$ **-On-Line-**

Define New Data of $\mathbf{P}(\ell+1)$

$$p_0 = \begin{bmatrix} \bar{x}_0^\ell \\ \boxed{y_\ell} \\ y_{\ell-1} \\ \vdots \\ y_{\ell-N} \end{bmatrix} \xleftarrow{\text{Horizon Shifting}} \xrightarrow{} p = \begin{bmatrix} \bar{x}_0^{\ell+1} \\ y_{\ell+1} \\ \boxed{y_\ell} \\ \vdots \\ y_{\ell-N+1} \end{bmatrix}$$

$$\hat{s}^*(p) \approx s^*(p_0) + \left.\frac{\partial s}{\partial p}^T\right|_{p_0} (p - p_0)$$

Fast Approximate Solution

Large Perturbation $||p - p_0||$ **- Might Give Poor Approximations**

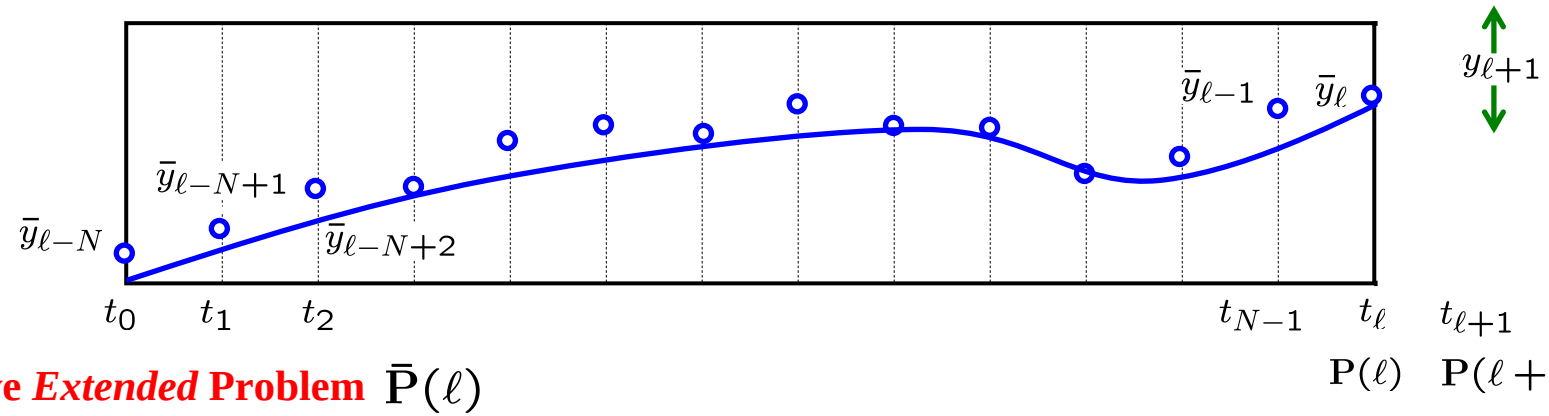
How to Consider Moving Horizon Shifting?

Fast Moving Horizon Estimation

How to Keep $||p - p_0||$ **Small ?**

- Shifted Approach

Free Dummy Variable



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

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

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

***Extended* Lagrangian**

$$\bar{\mathcal{L}} = \boxed{\mathcal{L}} + \boxed{L_{N+1}(v_{N+1}, w_{N+1}) + \lambda_{N+1}^T (x_{N+1} - f_N(x_N, w_N)) + \bar{\lambda}_{N+1}^T (y_{\ell+1} - h_{N+1}(x_{N+1}) - v_{N+1})}$$

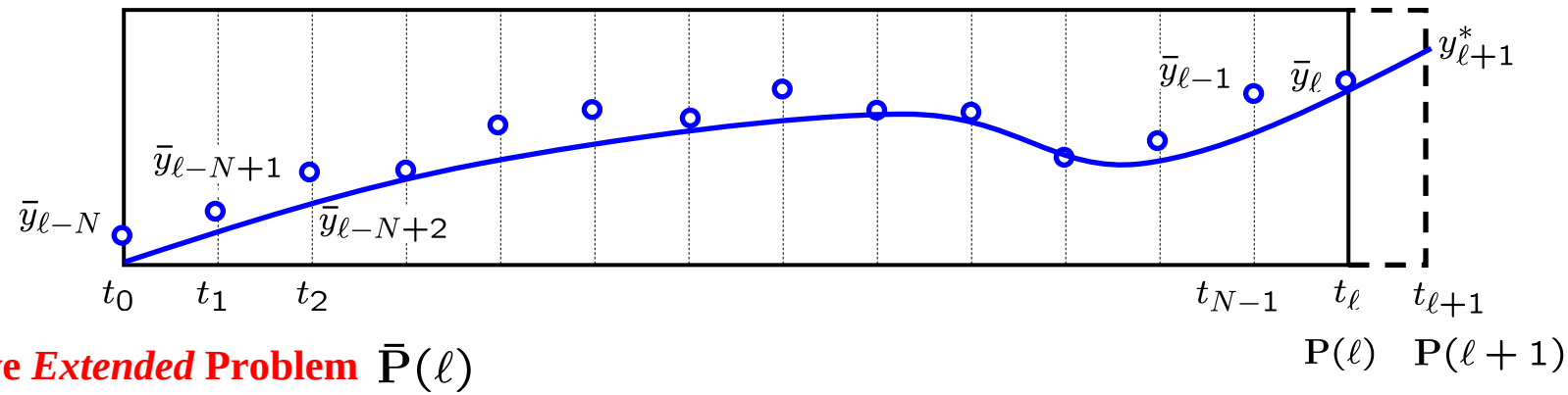
$\mathbf{P}(\ell)$

$$\nabla_{y_{\ell+1}} \bar{\mathcal{L}} = \bar{\lambda}_{N+1} = 0$$

$$\nabla_{v_{N+1}} \bar{\mathcal{L}} = \mathbf{R}_{N+1}^{-1} v_{N+1} - \bar{\lambda}_{N+1} = 0 \longrightarrow v_{N+1} = 0$$

Dummy Variable = Output Model Extrapolation

Fast Moving Horizon Estimation



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

KKT System At Solution of $\bar{\mathbf{P}}(\ell)$

$$\begin{bmatrix} H_k & A_k & -I \\ A_k^T & 0 & 0 \\ V_k & 0 & X_k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta \nu \end{bmatrix} = -\begin{bmatrix} \nabla f(x^*) + A_k \lambda^* - \nu^* \\ c(x^*) \\ X^* V^* e - \mu_\ell e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Factorization of KKT Matrix Available

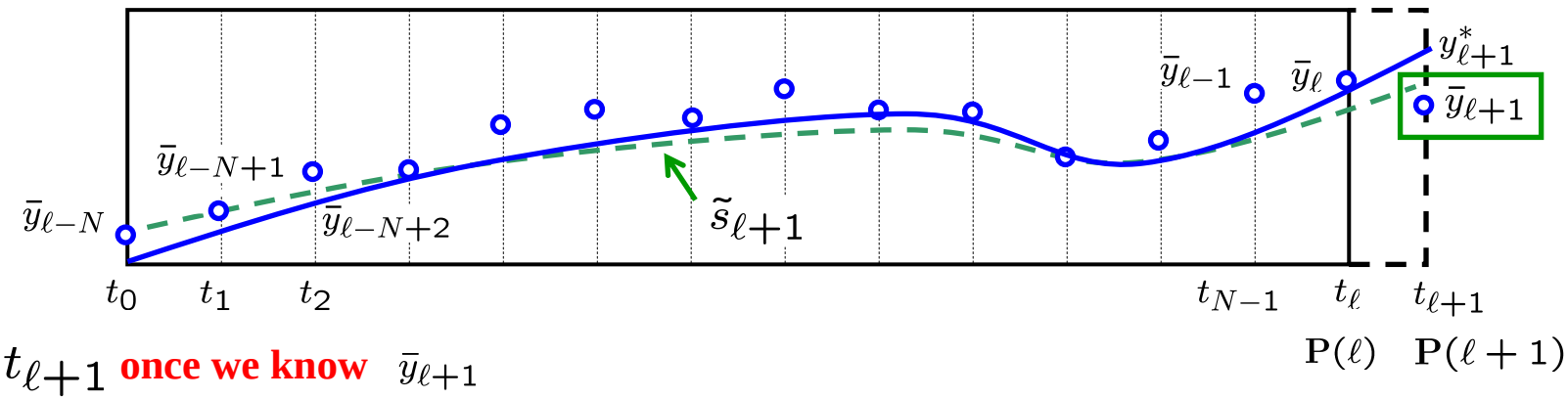
$$\mathbf{K}\Delta s = -\varphi(s^*) = 0$$

Extended Horizon Terms

$$\Delta \bar{\lambda}_{N+1} = 0$$

$$\mathbf{R}_{N+1}^{-1} \Delta v_{N+1} + \Delta \bar{\lambda}_{N+1} = 0$$

Fast Moving Horizon Estimation



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

KKT System of $\bar{\mathbf{P}}(\ell)$

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

$$\Delta\bar{\lambda}_{N+1} = 0$$

$$\mathbf{R}_{N+1}^{-1}\Delta v_{N+1} + \Delta\bar{\lambda}_{N+1} = 0$$

$$\Delta\bar{\lambda}_{N+1} + \boxed{\Delta p} = 0$$ **Relax Multiplier**

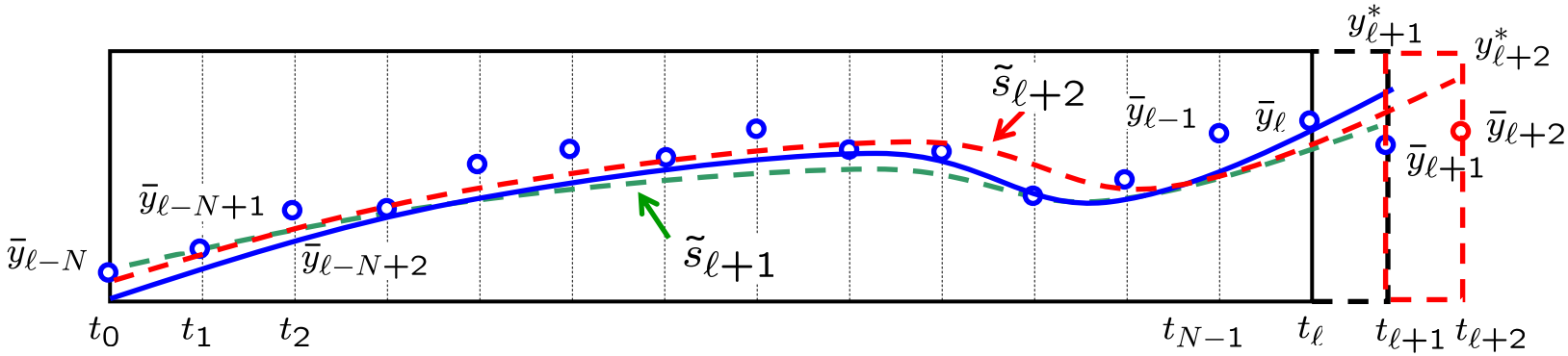
$$\boxed{\Delta y_{\ell+1} = (y^*_{\ell+1} - \bar{y}_{\ell+1})}$$ **Force Variable to Measurement**

Augmented KKT System

$$\tilde{s}_{\ell+1} = s^*_\ell + \Delta s \qquad \left[\begin{array}{c|c} \mathbf{K} & \mathbf{E_p} \\ \hline \mathbf{E_y}^T & 0 \end{array}\right] \left[\begin{array}{c} \Delta s \\ \hline \Delta p \end{array}\right] = -\left[\begin{array}{c} 0 \\ \hline \bar{y}_{\ell+1} - y^*_{\ell+1} \end{array}\right]$$

Fast On-line Solution?

Fast Moving Horizon Estimation



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

Schur Complement Technique

$$-(\mathbf{E_y}^T \mathbf{K}^{-1} \mathbf{E_p}) \Delta p = -(\bar{y}_{\ell+1} - y^*_{\ell+1})$$

- Background –
Form Schur Complement

Fast Approximate Solution

$$\tilde{s}_{\ell+1} = s^*_\ell + \Delta s \qquad \mathbf{K} \Delta s = -\mathbf{E_p} \Delta p$$

- On-Line –
Single Back Solve

→ **Update in Background Every Sampling Time**

Larger Models - Extend Horizon Further

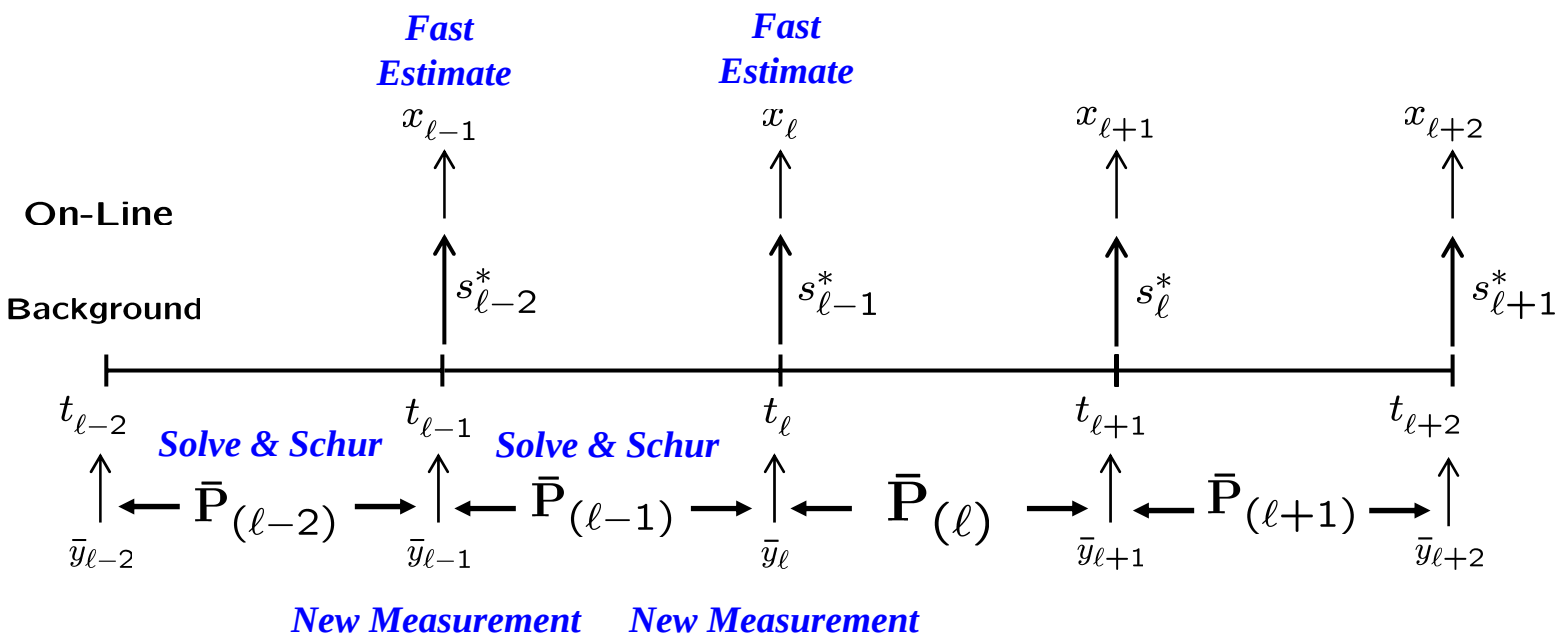
$$\tilde{s}_{\ell+2} = s^*_\ell + \Delta s \qquad \mathbf{K} \Delta s = [\mathbf{E_{p_1}}\ \mathbf{E_{p_2}}] \begin{bmatrix} \Delta p_1 \\ \Delta p_2 \end{bmatrix}$$

- On-line -
Single Back Solve

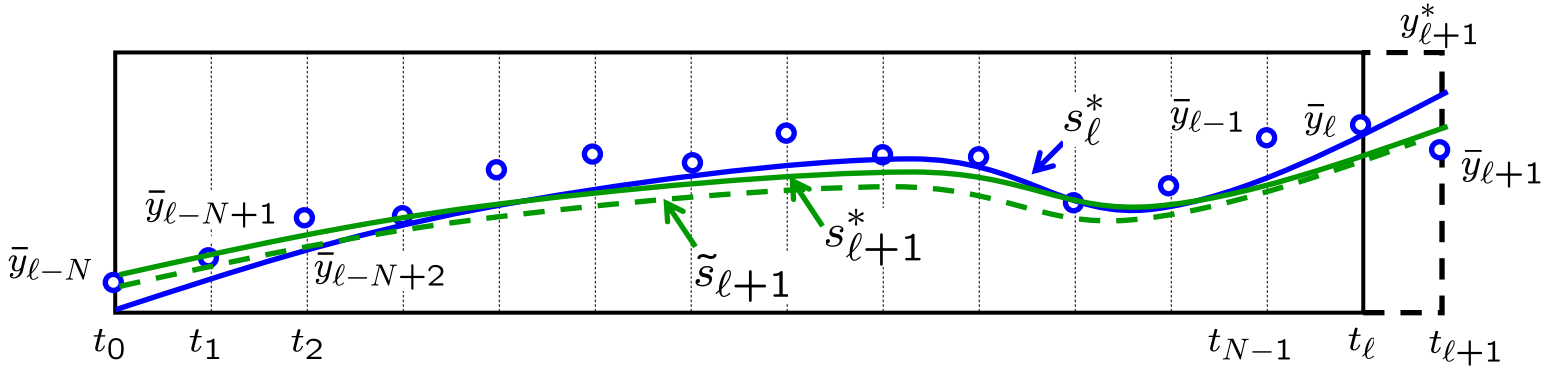
→ **Update in Background Every Two Sampling Times**

Fast Moving Horizon Estimation

Schedule of Computational Tasks



Fast Moving Horizon Estimation



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

$$-(\mathbf{E_y}^T \mathbf{K}^{-1} \mathbf{E_p}) \Delta p = -(\bar{y}_{\ell+1} - y^*_{\ell+1})$$

$$\mathbf{K} \Delta s = -\mathbf{E_p} \Delta p \qquad \tilde{s}_{\ell+1} = s^*_\ell + \Delta s$$

Approximation Error

$$\|\tilde{s}_{\ell+1} - s^*_{\ell+1}\| \leq L \|\bar{y}_{\ell+1} - y^*_{\ell+1}\|$$

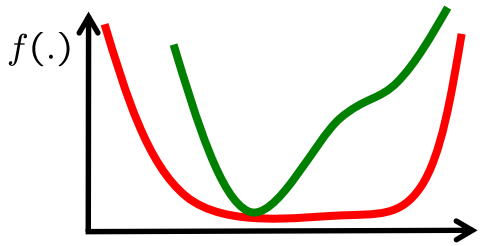
Solve $\mathbf{K} \Delta s = -\mathbf{E_p} \Delta p$ **through Forward Riccati Decomposition**

$$\|\tilde{x}_{\ell+1} - x^*_{\ell+1}\| \leq L_z \|\mathbf{\Pi}_{N+1}\| \times \|\bar{y}_{\ell+1} - y^*_{\ell+1}\|$$

Strongly Observable $\lambda_{min}(\mathbf{\Pi}_{N+1}) \to 0$ **-Approximation Error Small-**

Weakely Observable $\lambda_{min}(\mathbf{\Pi}_{N+1}) \to \infty$ **-Approximation Error Large-**

Strongly Observable and $N \to \infty$ $\|\tilde{x}_{\ell+1} - x^*_{\ell+1}\| \to 0$



Numerical Analysis of $\mathbf{\Pi}_{N+1}$

Industrial Case Study

Industrial Case Study

Low-Density Polyethylene Tubular Reactor Process

Material Balances - ODEs

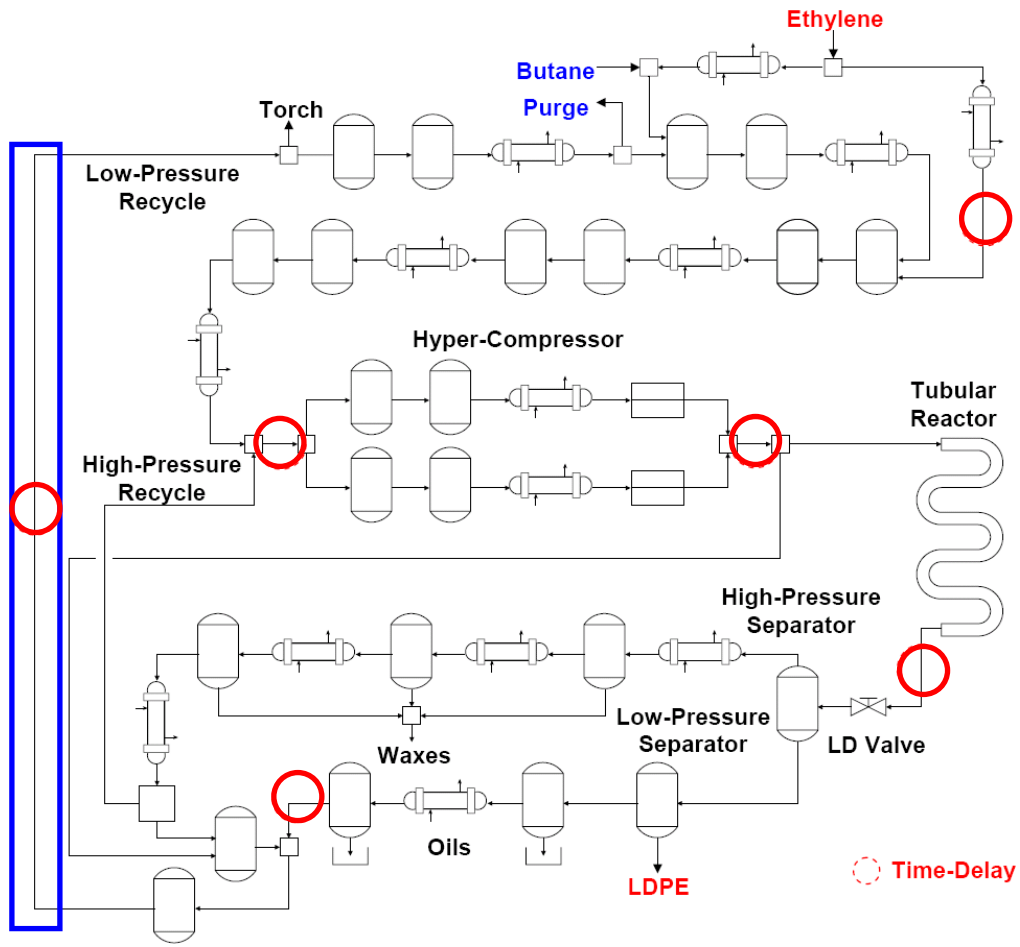
$$\frac{d\left(V_k \rho_k w_k^j\right)}{dt} = F_k\, w_{IN_k}^j - F_k w_k^j$$
$$k \in \{1..N_U\}$$
$$j \in \{1..N_C\}$$

Variable Time Delays - PDEs

$$\frac{\partial w_i^j}{\partial t} + \frac{1}{\tau}\frac{\partial w_i^j}{\partial z} = 0$$
$$w_i^j(z,0) = w_{0_i}^j$$
$$i \in \{1..N_T\}$$
$$j \in \{1..N_C\}$$

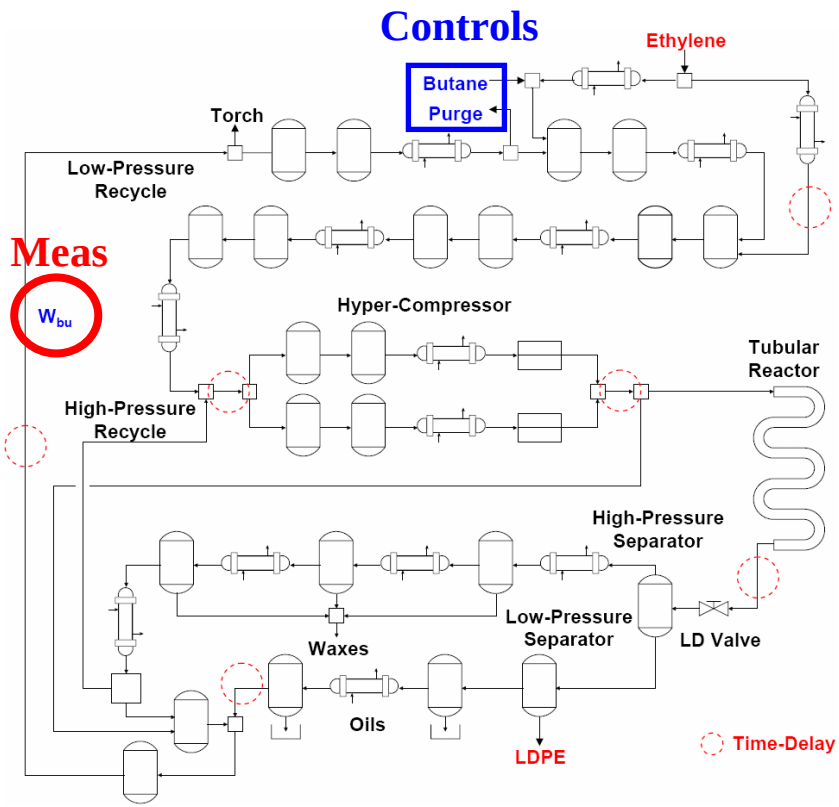
Discretize
-in Space-

Dynamic Process Model
294 ODEs
500 AEs



High-Pressures – Lots of Compression Stages
Low Reactor Conversions
Long Recycle Loops – Variable Time Delays

Industrial Case Study



Single Measurement - Composition of Recycle Gas
Sampling Time = 6 min
Estimation Horizon N = 15

NLP Simultaneous Approach

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

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

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

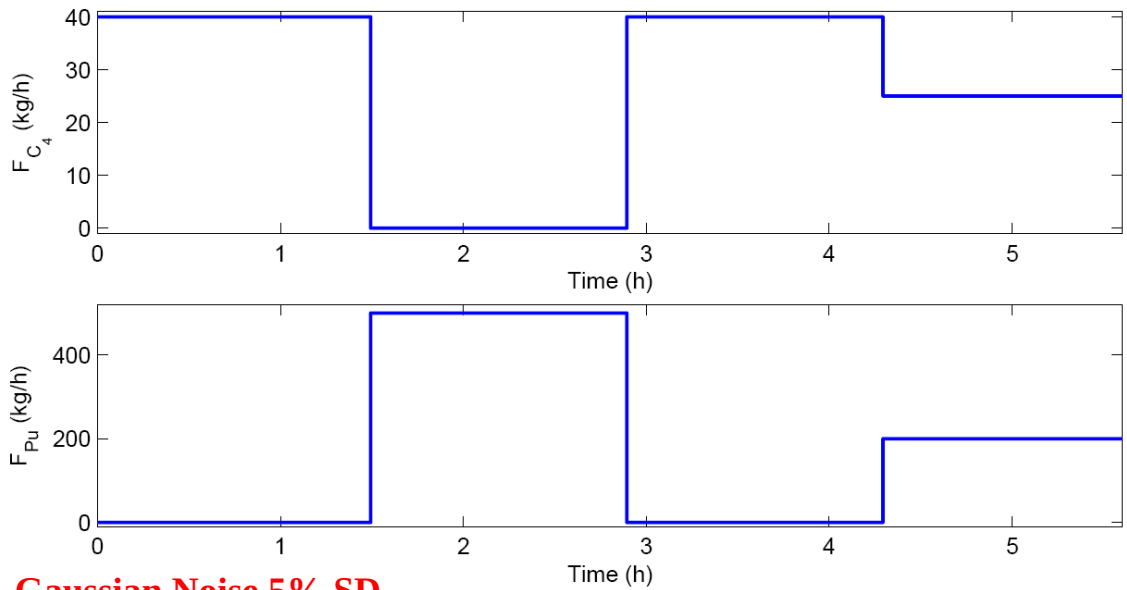
← **On-Line Cost Standard MHE**

← **On-Line Cost Fast MHE**

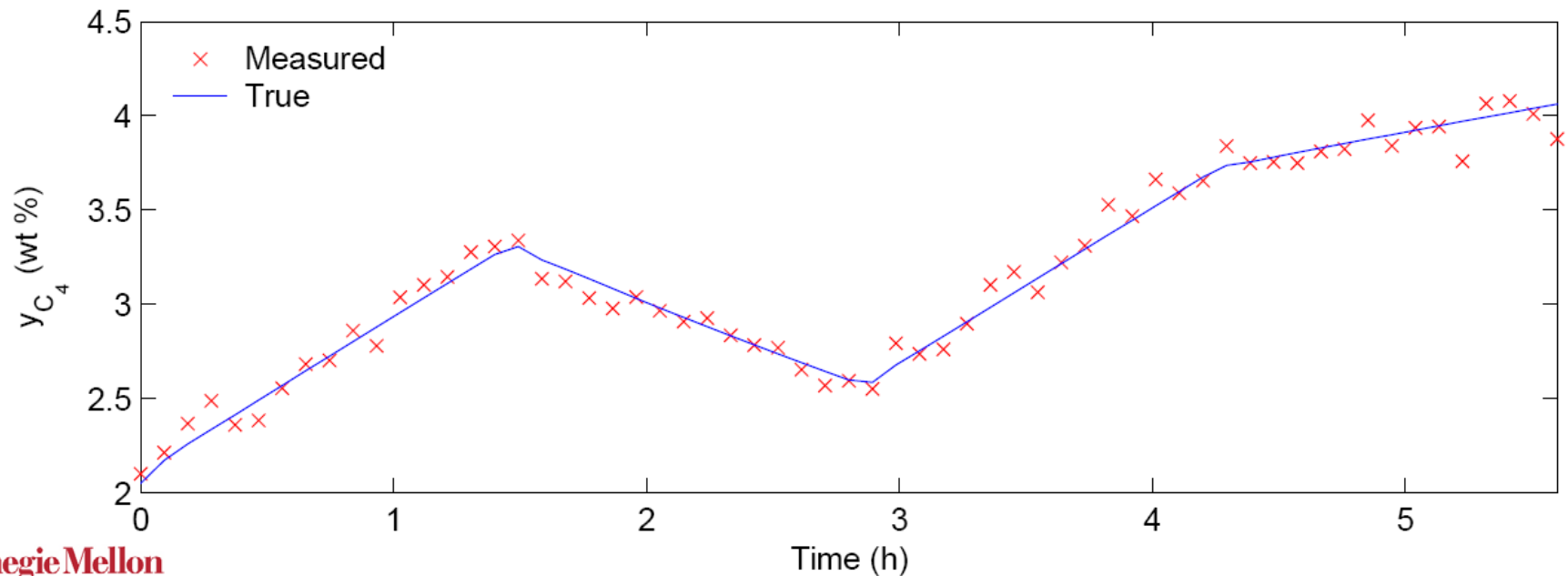
Solutions Satisfy Second Order Conditions
– System Locally Observable -

Industrial Case Study

Simulated Control Profiles

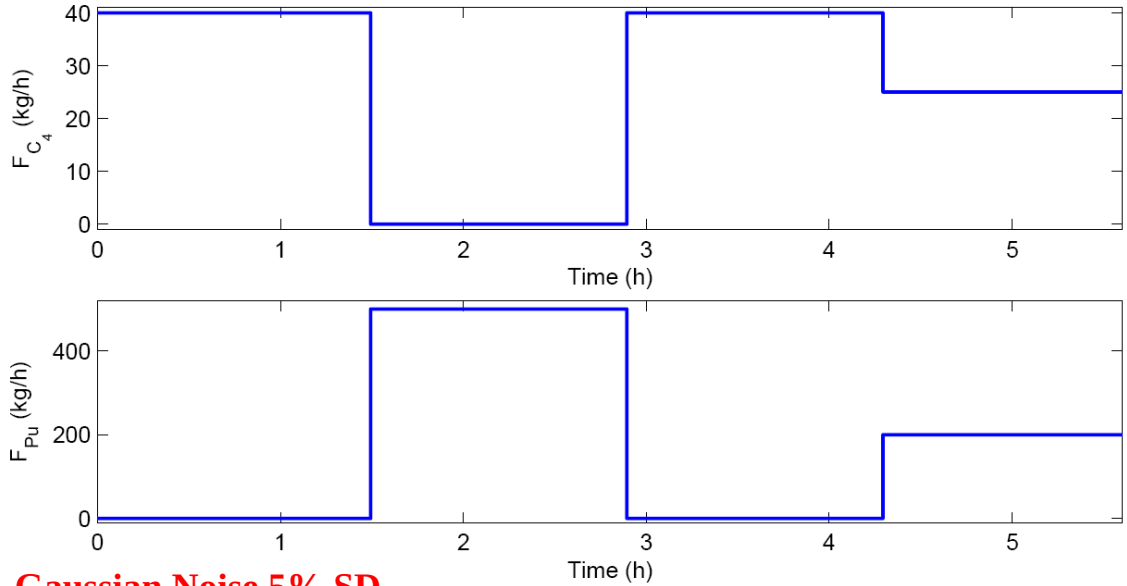


Measurement Profiles Gaussian Noise 5% SD

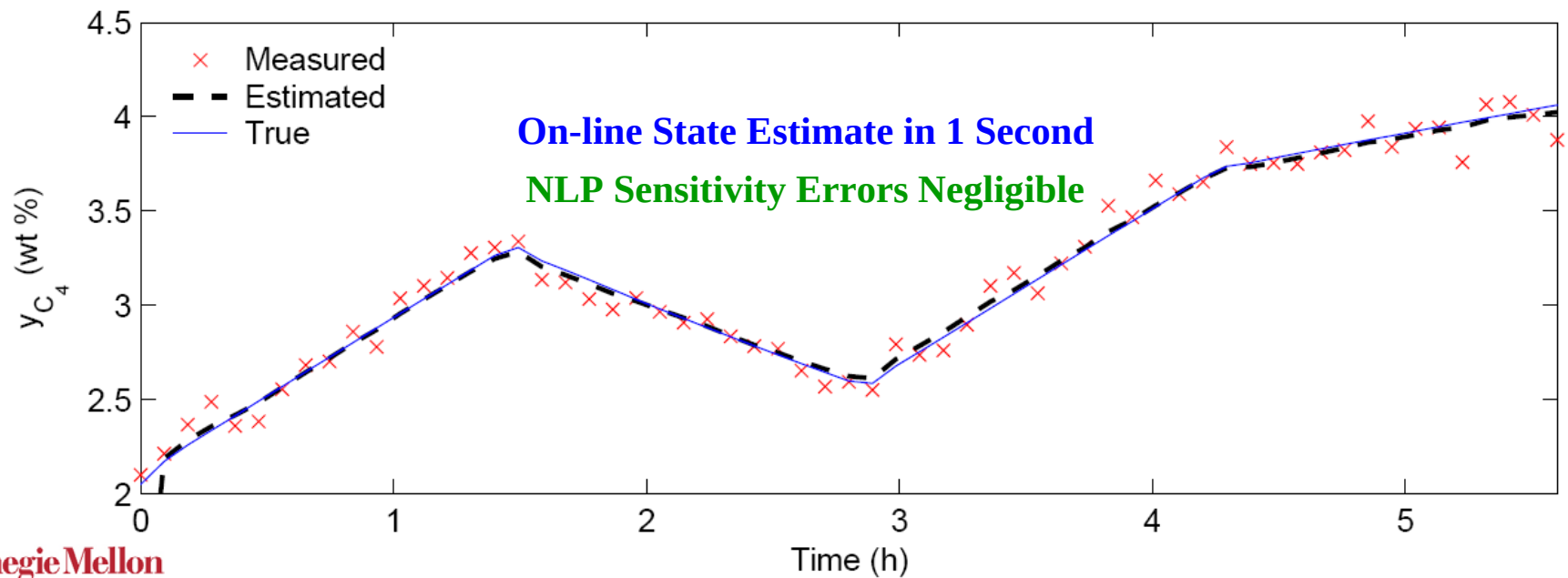


Industrial Case Study

Simulated Control Profiles

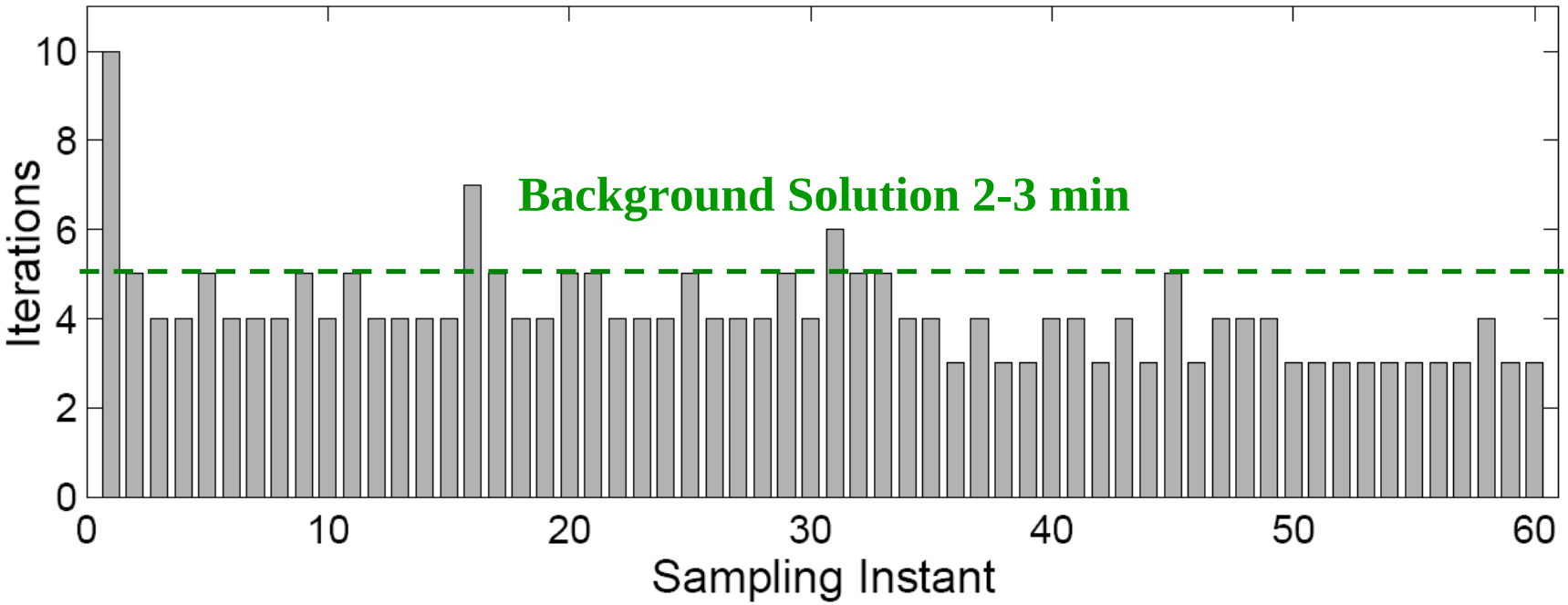


Measurement Profiles Gaussian Noise 5% SD



Industrial Case Study

Performance of NLP Solver In Background



Fast MHE = Fast Accurate Approximation + Fast Solver in Background

Conclusions and Future Work

Conclusions

Fast Computational Framework for MHE

- **Simultaneous Collocation Approach -Computational Complexity & Exact Derivatives**
- **Reference Problem in Background**
- **Fast On-Line Approximations – NLP Sensitivity**
- **Exploit Parametric Property with Appropriate Formulations**
- **Simulated Industrial Scenario - Accurate State Estimates**
  - Reduction of On-Line Costs by Factor of 200

Future Work

Stability of Fast MHE Algorithm

Lyapunov-Based Analysis *Rao, Mayne & Rawlings 2003*– **NLP Sensitivity Errors** *Z. & Biegler 2007*

Fast Covariance Information from Large-Scale NLP Solver

Full-Space KKT Matrix - EKF Riccati Recursion *Z. & Biegler, 2007*

AIChE Annual Meeting 2007

Complete Fast NMPC Framework

Fast NMPC *Z., Laird & Biegler 2006-2007* **+ Fast MHE** *This Work*

Faster MHE …

Alternative Formulations, Parallel Decomposition, Linear Solvers

Acknowledgments

Dr. Don Bartusiak
Dr. Andreas Waechter
Prof. Dr. Moritz Diehl and Heidelberg Group

A Fast Computational Framework for Large-Scale Moving Horizon Estimation

Victor M. Zavala, Carl D. Laird and Lorenz T. Biegler

Department of Chemical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

vzavala@cmu.edu

http://dynopt.cheme.cmu.edu

8th International Symposium on Dynamics and Control of Process Systems
June 8th 2007

Carnegie Mellon

CAPD CENTER