Real-Time Strategies for Control Architectures using Differential Variational Inequalities

Victor M. Zavala and Mihai Anitescu

Mathematics and Computer Science Division
Argonne National Laboratory
vzavala@mcs.anl.gov

Midwest Numerical Analysis Conference
April 24th, 2010

Outline of the Talk

Motivation

- **Interplay System Dynamics and Solution Accuracy**

Real-Time Optimization and Feedback

- **Stability and Approximation Errors**
- **Linear Algebra**
- **Generalized Equations**
- **Augmented Lagrangian**

Future Work

- **Real-Time Grid Simulation**

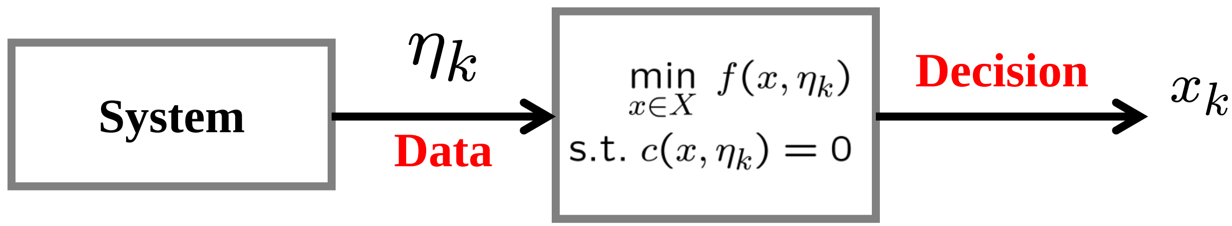
Motivation

Motivation

Optimization Domains:

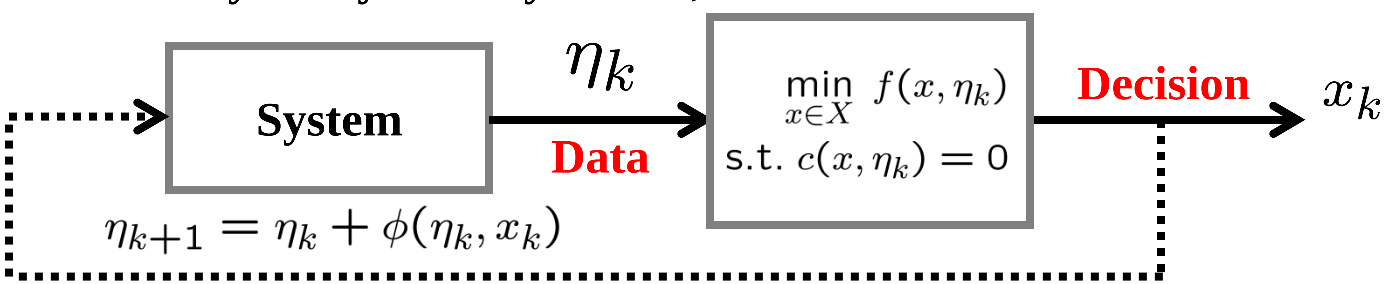
Design: - Single Data Instance

- Solution Time Not Critical, High Accuracy, No Warm-Start

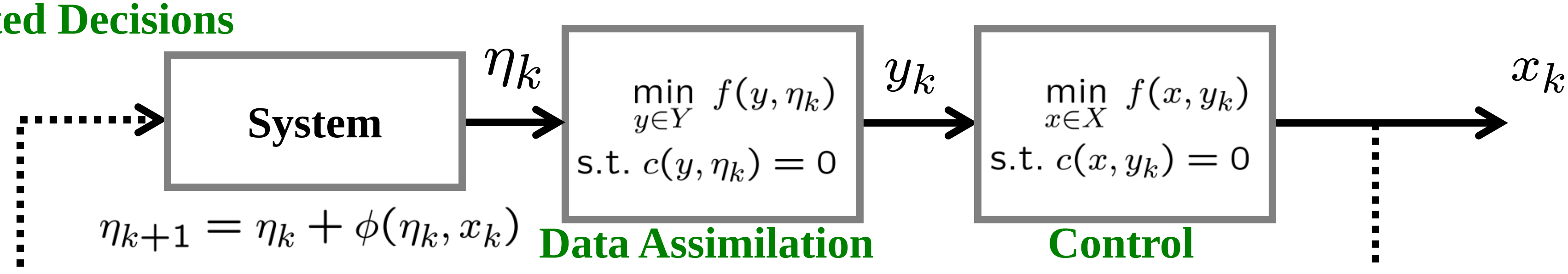


Operational: - Real-Time Data Flow

- Performance Affected by Solution Frequency and Delay
- Accuracy vs. System Dynamics, Warm-Start



Nested Decisions



Motivation

Applications:

- **Data Assimilation, Model Predictive Control, Dynamic Games**
- **Weather Forecasting, Power Flow Control, Buildings Control, Energy Management, …**

$$
\begin{aligned}
\min_{u(\tau),z(t)} \quad & \int_t^{t+T} \varphi\left(z(\tau), y(\tau), u(\tau), \eta(t)\right) d\tau \\
\text{s.t.} \; \frac{dz}{d\tau} &= \mathrm{f}(z(\tau), y(\tau), u(\tau), \eta(t)) \\
0 &= \mathrm{g}(z(\tau), y(\tau), u(\tau), \eta(t)) \\
0 &\geq \mathrm{h}(z(\tau), y(\tau), u(\tau), \eta(t))
\end{aligned}
$$

DAE-Constrained Optimization

Finite Elements →

$$
\begin{aligned}
&\min_{x \in X} \; f(x, \eta(t)) \\
&\text{s.t.} \; c(x, \eta(t)) = 0
\end{aligned}
$$

NLP

Trend is to Use Detailed Physical Models ~10,000 DAEs in 2 Minutes *Z & Biegler, 2008*

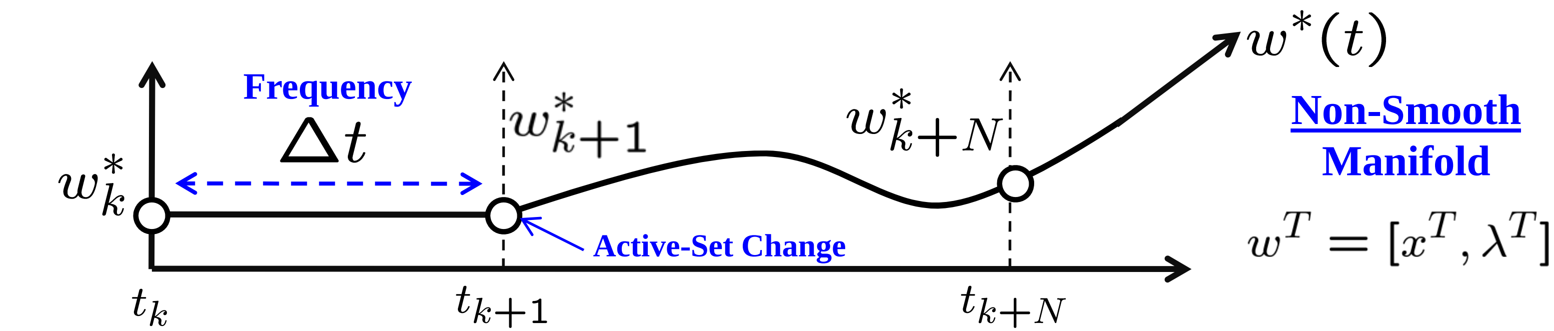
Physical Models Are <u>Difficult to Decompose</u> -Limits Parallelism-

How Fast Can We and Do We Need to Solve?

Model Complexity vs. System Dynamics

Real-Time Optimization and Feedback

Real-Time Optimization



$$\min_{x \in X} f(x, \eta_k) \quad \text{s.t. } c(x, \eta_k) = 0$$

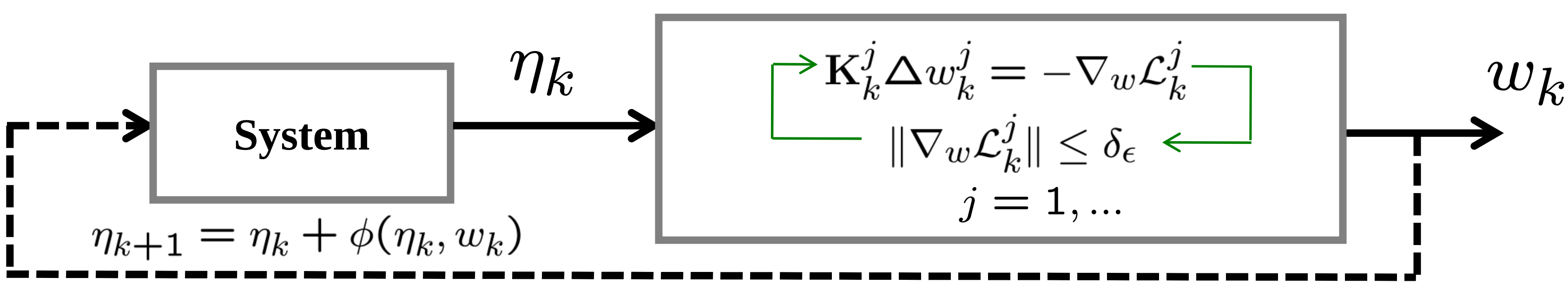
$$\min_{x \in X} f(x, \eta_{k+1}) \quad \text{s.t. } c(x, \eta_{k+1}) = 0$$

$$\min_{x \in X} f(x, \eta_{k+N}) \quad \text{s.t. } c(x, \eta_{k+N}) = 0$$

Problems Get Closer to Each Other as $\Delta t \rightarrow 0$ -Solution Time-

The "Easy" Case – No Inequalities – Smooth Manifold

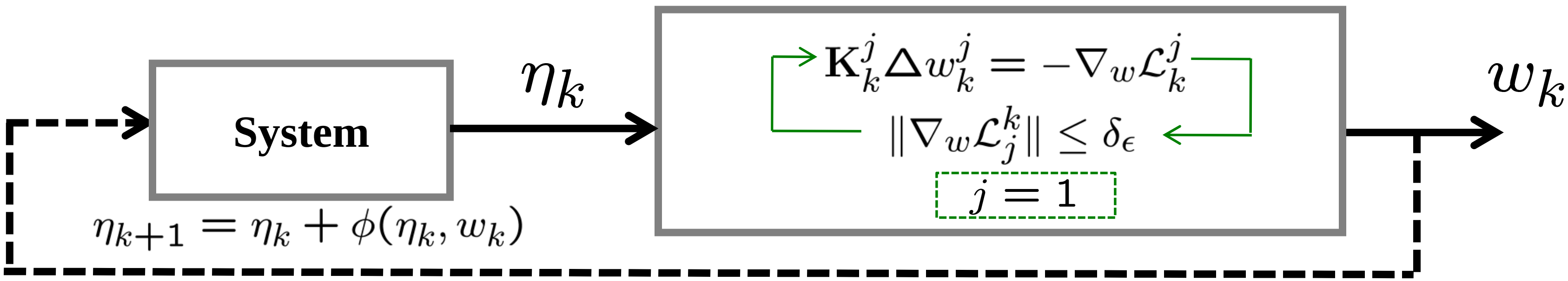
$$\min_{x \in \Re} f(x, \eta(t)) \quad \text{s.t. } c(x, \eta(t)) = 0 \longrightarrow \nabla_w \mathcal{L}(w, \eta(t)) = 0$$



High Accuracy Brings $w(t)$ To Manifold $w^*(t)$ But Increases Δt

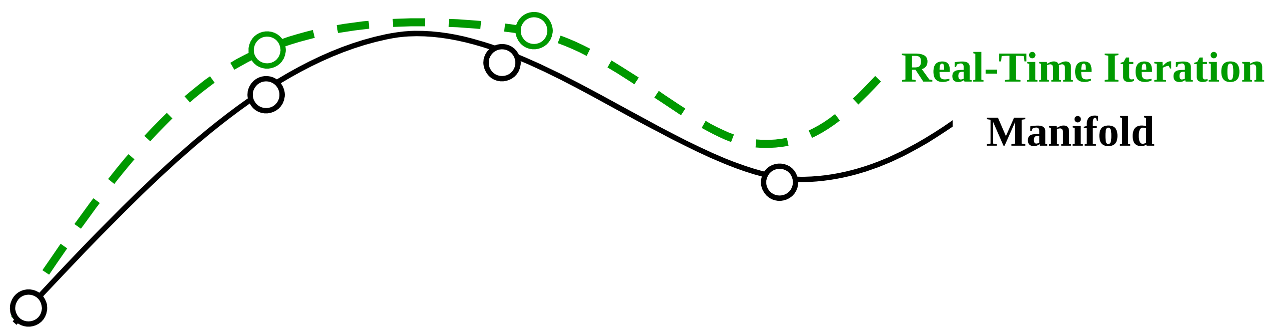
Stability and Approximation Errors

"Real-Time Iteration" – Fix Computational Time *Diehl, 2001*



Stability: If $\|\Delta \eta_k\| \leq \delta_\eta$ **and** $\|\nabla_w \mathcal{L}_{j,i}^k\| \leq \delta_\epsilon$ **Then** $\|x_k - x_k^*\| \leq \delta_x, \;\; k > 0$

Convergence: If $\|\Delta \eta_k\| \to 0$ **Then** $x_k \to x_k^*, \;\; k > 0$



First Iteration is Tangential Predictor, Subsequent Iterations are Corrector Steps *Jongen, 1996*

Interpretation as a Differential Equation *Ohtsuka, 2004* *Z & Anitescu,2009*

$$\frac{d\nabla_w \mathcal{L}(w(t), \eta(t))}{dt} = 0 \longrightarrow \begin{array}{c} \mathbf{K}(t)\dot{w} = -\nabla_{w,\eta}\mathcal{L}(t)\,\dot{\eta} \\ \dot{\eta} = \phi(\eta(t), w(t)) \end{array} \longrightarrow \begin{array}{c} \mathbf{K}_k \Delta w_k = -\nabla_{w,\eta}\mathcal{L}_k \Delta\eta_k \\ \Delta\eta_k = \phi(\eta_k, w_k) \end{array}$$

Explicit Euler

Linear Algebra

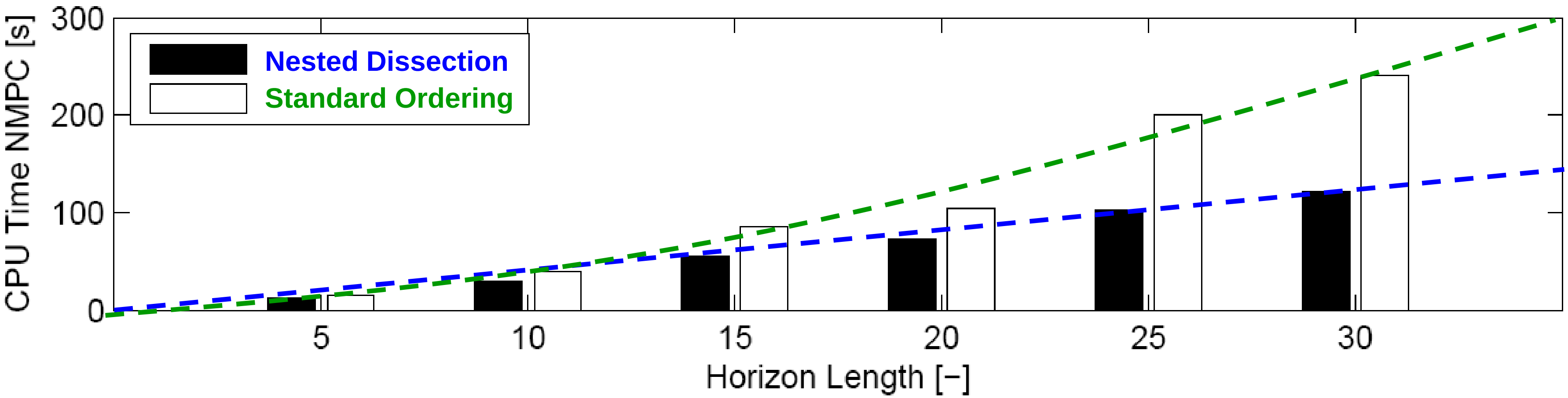
Newton Step Computation

Symmetric Indefinite System

$$\begin{bmatrix} \mathbf{H}_k^j & \mathbf{A}_k^{j\,T} \\ \mathbf{A}_k^j & \end{bmatrix} \begin{bmatrix} \Delta x_k^j \\ \Delta \lambda_k^j \end{bmatrix} = \begin{bmatrix} \nabla_x \mathcal{L}(w_k^j, \eta_k) \\ \nabla_\lambda \mathcal{L}(w_k^j, \eta_k) \end{bmatrix}$$

Direct Factorizations (MA57, Pardiso) – Complexity $O(n_w^{1,2})$

Performance of MA57 in Model Predictive Control *Z & Biegler, 2008*

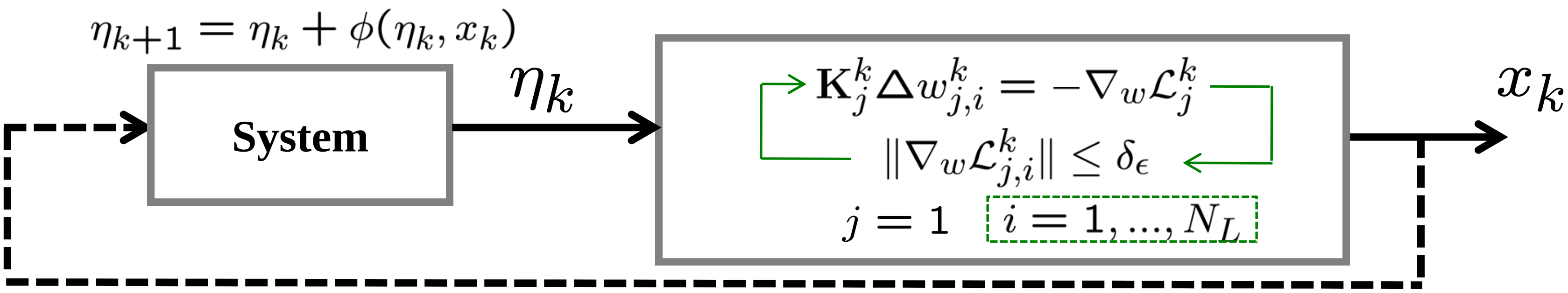


<u>One</u> Iteration for NLP with <u>350,000</u> Variables and Constraints ~ <u>20 CPUs</u>

Direct Solvers Cannot Be Warm-Started, Memory Bottlenecks with $n_w = O(10^6)$

Linear Algebra

<u>Truncated</u> Matrix-Free Newton (PCG, QMR,GMRES)?



Stability Still Preserved as Long as $\|\nabla_w \mathcal{L}_{j,N_L}^k\| \leq \delta_\epsilon$ **- Terminating Early** $\Delta t \to 0$ **-**

How to Deal with <u>Inequalities</u>?

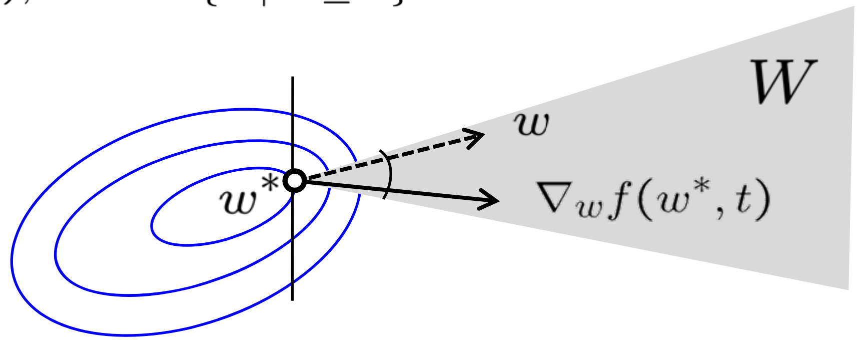
Generalized Equations

Generalized Equations (GE) *Robinson, 1977, 1980*

$F(w,t) \in \mathcal{N}_W(w)$ ⟵ **Normal Cone Operator**

First-Order KKT Conditions of $\min_{w \in W} \; f(w,t), \quad W = \{w \,|\, w \geq 0\}$

$\nabla_w f(w^*,t)^T(w - w^*) \geq 0, \; \forall w \in W$



Effect of Perturbations - Linearized Generalized Equation (LGE)

$\delta \in F(w_0^*, t_0) + \nabla_w F(w_0^*, t_0)(w - w_0^*) + \mathcal{N}_W(w) \qquad w(\delta) = \psi^{-1}[\delta]$ ⟵ **Solution Operator**

Definition: **LGE is Strongly Regular at** w_0^* **if** $\exists\, L_\psi \geq 0 \text{ s.t. } \|w(\delta) - w_0^*\| \leq L_\psi \|\delta\|$

$$\left.\begin{array}{l} \min_{x \in X} \; f(x,t) \\ \text{s.t. } c(x,t) = 0 \end{array}\right\}$$

$\nabla_w \mathcal{L}(w,t) \in \mathcal{N}_{X \times \Re^m}(w)$ **GE**

$\delta \in \nabla_w \mathcal{L}(w_0^*, t_0) + \mathbf{K}(w_0^*, t_0)(w - w_0^*) + \mathcal{N}_{X \times \Re^m}(w)$ **LGE**

Theorem: ψ^{-1} **is Lipschitzian if SSOC and LICQ Hold at** w_0^*

Generalized Equations

Perturbed LGE w_0^*

$$\delta \in \nabla_w \mathcal{L}(w_0^*, t) + \mathbf{K}(w_0^*, t_0)(w - w_0^*) + \mathcal{N}_{X \times \Re^m}(w)$$

Canonical Form $\Updownarrow$

$$\delta \in F(w_0^*, t_0) + \nabla_w F(w_0^*, t_0)(w - w_0^*) + \mathcal{N}_W(w)$$

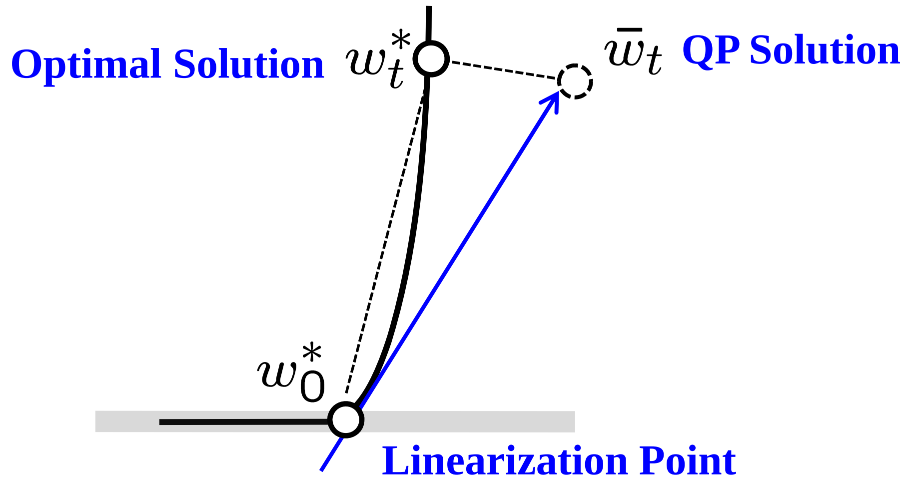
Perturbed LGE = Perturbed QP

$$\begin{aligned} \min \quad & \nabla_x f(x_0^*, t)^T \Delta x + \frac{1}{2} \Delta x^T \mathbf{H}_0^* \Delta x \\ \text{s.t.} \quad & c(x_{t_0}^*, t) + \mathbf{A}_0^* \Delta x = 0 \\ & \Delta x \geq -x_{t_0}^* \end{aligned}$$

$$\delta = \nabla_w \mathcal{L}(w_0^*, t_0) - \nabla_w \mathcal{L}(w_0^*, t)$$

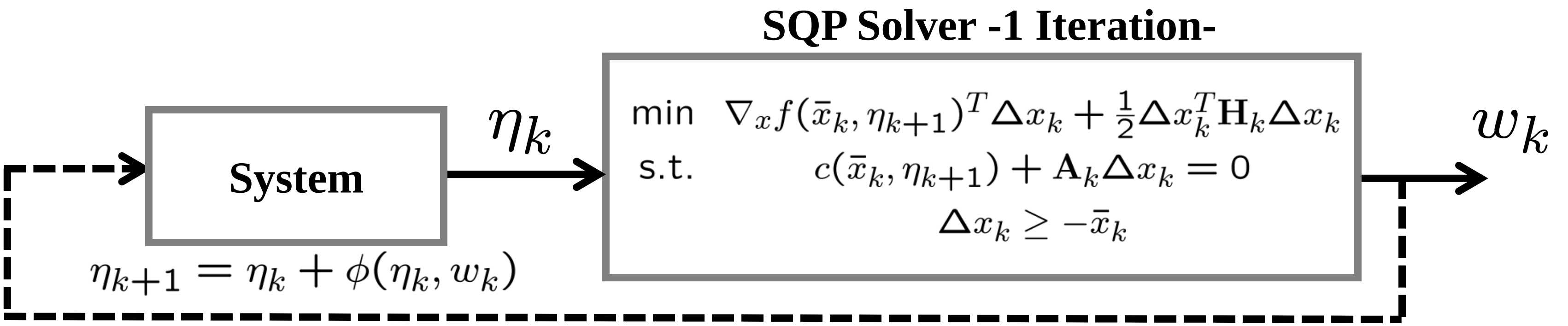
From Lipschitz Continuity and Mean Value Theorem (Non-Smoothness Do Not Show Up in Analysis)

$$\begin{aligned} \|w_t^* - \bar{w}_t\| & \leq L_\psi \|r(w_t^*, t) - \delta\| \\ & \leq L \Delta t^2 \end{aligned}$$



Key Observation: Error Bound Holds Even if Active-Set Changes

Generalized Equations



Theorem *Z & Anitescu, 2009*

- **A1: SSOC and LICQ Hold at** w_k^*
- **A2:** $\bar{w}_k$ **Exists in Neighborhood and** $\exists\, \delta_r \geq 0 \text{ s.t. } \|\bar{w}_k - w_k^*\| \leq L_\psi \|r(\bar{w}_k, \eta_k)\| \leq L_\psi \delta_r$

For Sufficiently Small Δt

$$\|\bar{w}_k - w_k^*\| \leq L_\psi \delta_r \Rightarrow \|\bar{w}_{k+1} - w_{k+1}^*\| \leq L_\psi \delta_r$$

Stability Holds Even QP Solution Error is $O(\Delta t^2)$

Interpretation as a Differential Variational Inequality (DVI)

Take $\Delta t \to 0$ **Then GE is a DVI** *Z & Anitescu, 2009*

Recursive QP Solution is Equivalent to Time-Stepping for DVI *Pang & Stewart, 2009*

Augmented Lagrangean

Augmented Lagrangean Penalty

$$\begin{aligned} \min_x \quad & f(x,\eta) \\ \text{s.t.} \quad & c(x,\eta) = 0 \\ & x \geq 0 \end{aligned} \quad \longleftrightarrow \quad \begin{aligned} \min \quad & \mathcal{L}_A(x,\lambda,\eta) := f(x,\eta) + \lambda^T c(x,\eta) + \frac{\rho}{2}\|c(x,\eta)\|^2 \\ \text{s.t.} \quad & x \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & \nabla_x \mathcal{L}_A(\bar{x}_k, \bar{\lambda}_k, \eta_{k+1})^T \Delta x_k + \tfrac{1}{2}\Delta x_k^T \nabla_{xx}\mathcal{L}_A(\bar{x}_k, \bar{\lambda}_k, \eta_k)\Delta x_k \\ \text{s.t.} \quad & \Delta x_k \geq -\bar{x}_k \end{aligned}$$

Close to Manifold **Hessian of AL Remains at Least Positive Semi-Definite** *Nocedal & Wright, 1999*

Projected Gauss Seidel

$$\min_{w \geq \alpha} \quad \frac{1}{2} w^T \mathrm{M} w + \mathrm{b}^T w$$

For $k = 0, 1, \ldots, N_L$

$$\begin{aligned} w_i^{k+1} &= -\frac{1}{\mathbf{M}_{ii}}\left(\mathbf{b}_i - \sum_{j<i} \mathbf{M}_{ij} w_j^{k+1} - \sum_{j>i} \mathbf{M}_{ij} w_j^k\right) \\ w_i^{k+1} &= \max\left(w_i^{k+1}, \alpha_i\right), \quad i = 1, \ldots, n \end{aligned}$$

- **Detects Multiple Active-Set Changes Efficiently** *Morales et.al. 2008, Tasora et.al. 2009*
- **High Accuracy Requires Large Number of Iterations → Less Important If Δt Small**

Augmented Lagrangean

Algorithm:

Given $\bar{x}_0, \bar{\lambda}_0$, Δt, ρ, and N_L,

1. Evaluate $\nabla_x \mathcal{L}_A(\bar{x}_k, \bar{\lambda}_k, \eta_{k+1}, \rho)$ and $\nabla_{xx} \mathcal{L}_A(\bar{x}_k, \bar{\lambda}_k, \eta_k, \rho)$.
2. Compute $\Delta \bar{x}_{k+1}$ applying N_L iterations to QP
3. Update $\bar{x}_{k+1} \leftarrow \bar{x}_k + \Delta x_k$ and $\bar{\lambda}_{k+1} \leftarrow \bar{\lambda}_k + \rho\, c(\bar{x}_{k+1}, \eta_{k+1})$.
4. $k \leftarrow k+1$

First-Order Multiplier Update, *Hestenes 1969*

AugLag Penalty Acts as Parametric Perturbation of Lagrange Multipliers, *Bertsekas, 1978*

Theorem *Z&Anitescu, 2009*

- **A1: Augmented Lagrangean - LGE is Strongly Regular at** w_k^*
- **A2:** $\bar{w}_k$ **Exists in Neighborhood and** $\exists\, \delta_r \geq 0$ s.t. $\|\bar{w}_k - w_k^*\| \leq L_\psi \|r(\bar{w}_k, \eta_k)\| \leq L_\psi \delta_r$

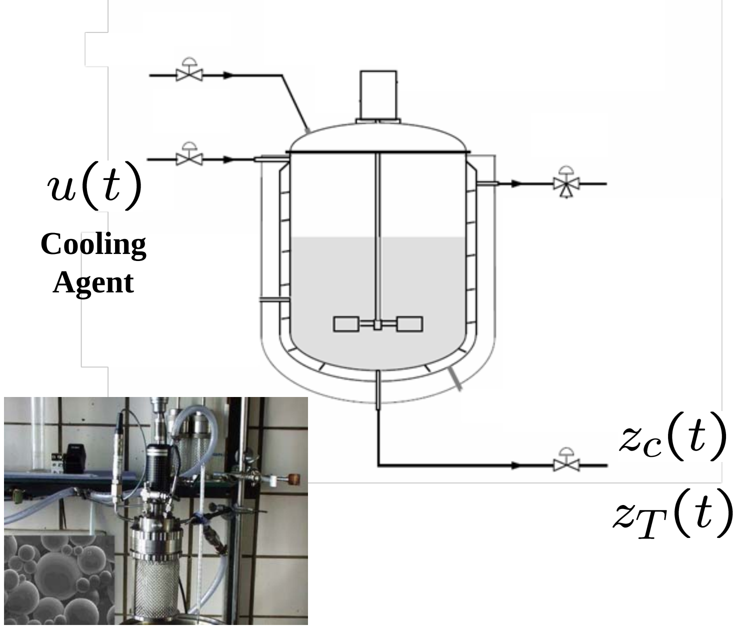
For sufficiently small Δt **and sufficiently large** ρ**,**

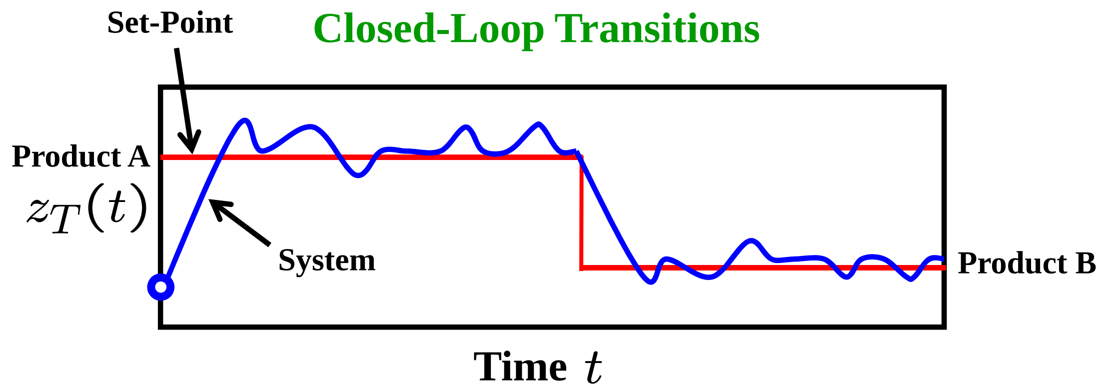
$$\|\bar{w}_k - w_k^*\| \leq L_\psi \delta_r \Rightarrow \|\bar{w}_{k+1} - w_{k+1}^*\| \leq L_\psi \delta_r$$

- **Conditions More Strict Due to Multiplier Error**
- **Tune** N_L **to Keep QP Solution Error** $O(\Delta t^2)$

Numerical Case Study

Control of Chemical Reactor





Optimal Control Problem

Set-Point

Time-Dependent Parameters

$$\min_{u(\tau)} \int_t^{t+T} \Big(w_T (z_T - z_T^{sp})^2 + w_C (z_C - z_C^{sp})^2 + w_u (u - u^{sp})^2 \Big)\, d\tau$$

$$\text{s.t.}\ \frac{dz_C}{d\tau} = \frac{z_C - 1}{\theta} + k_0 z_C \exp\left[\frac{-E_a}{z_T}\right], \qquad z_C(0) = z_C(t)$$

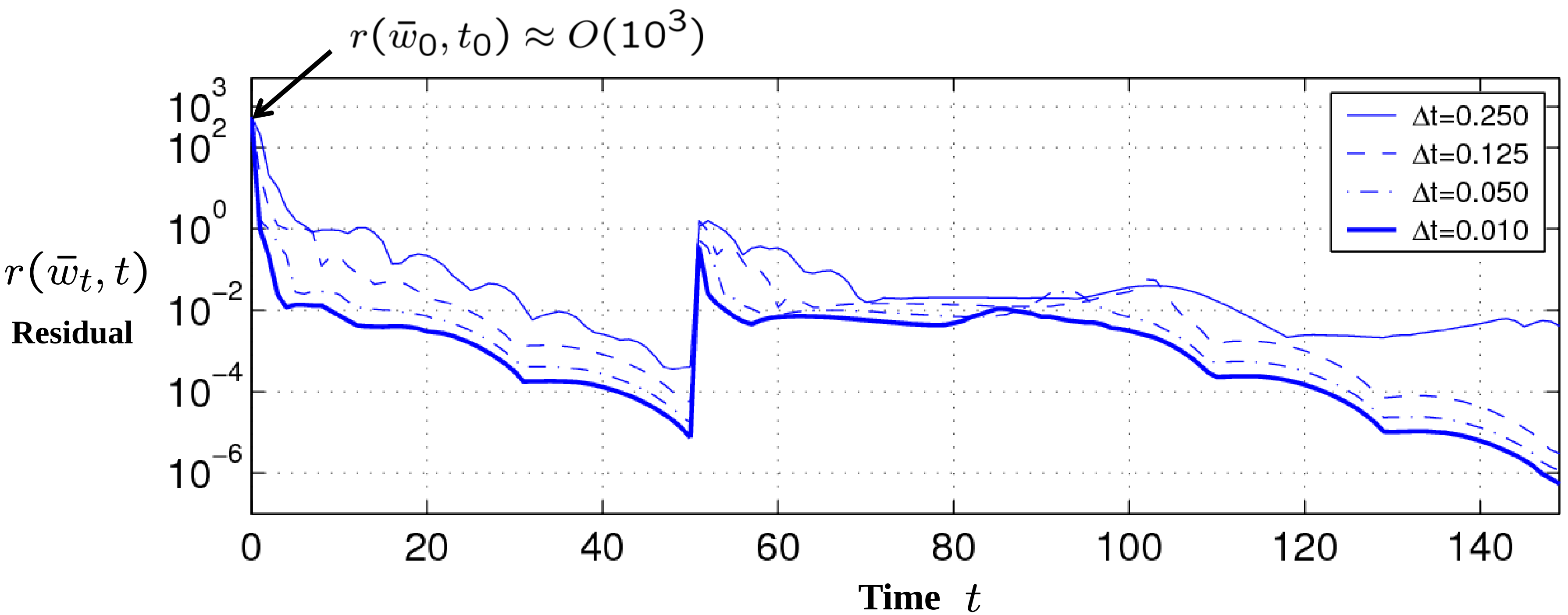
$$\frac{dz_T}{d\tau} = \frac{z_T - z_T^f}{\theta} - k_0 z_C \exp\left[\frac{-E_a}{z_T}\right] + \alpha u (z_T - z_T^{cw}), \qquad z_T(0) = z_T(t)$$

$$z_C^{min} \le z_C \le z_C^{max}, \quad z_T^{min} \le z_T \le z_T^{max}, \quad u^{min} \le u \le u^{max}.$$

Numerical Case Study

Numerical Tests

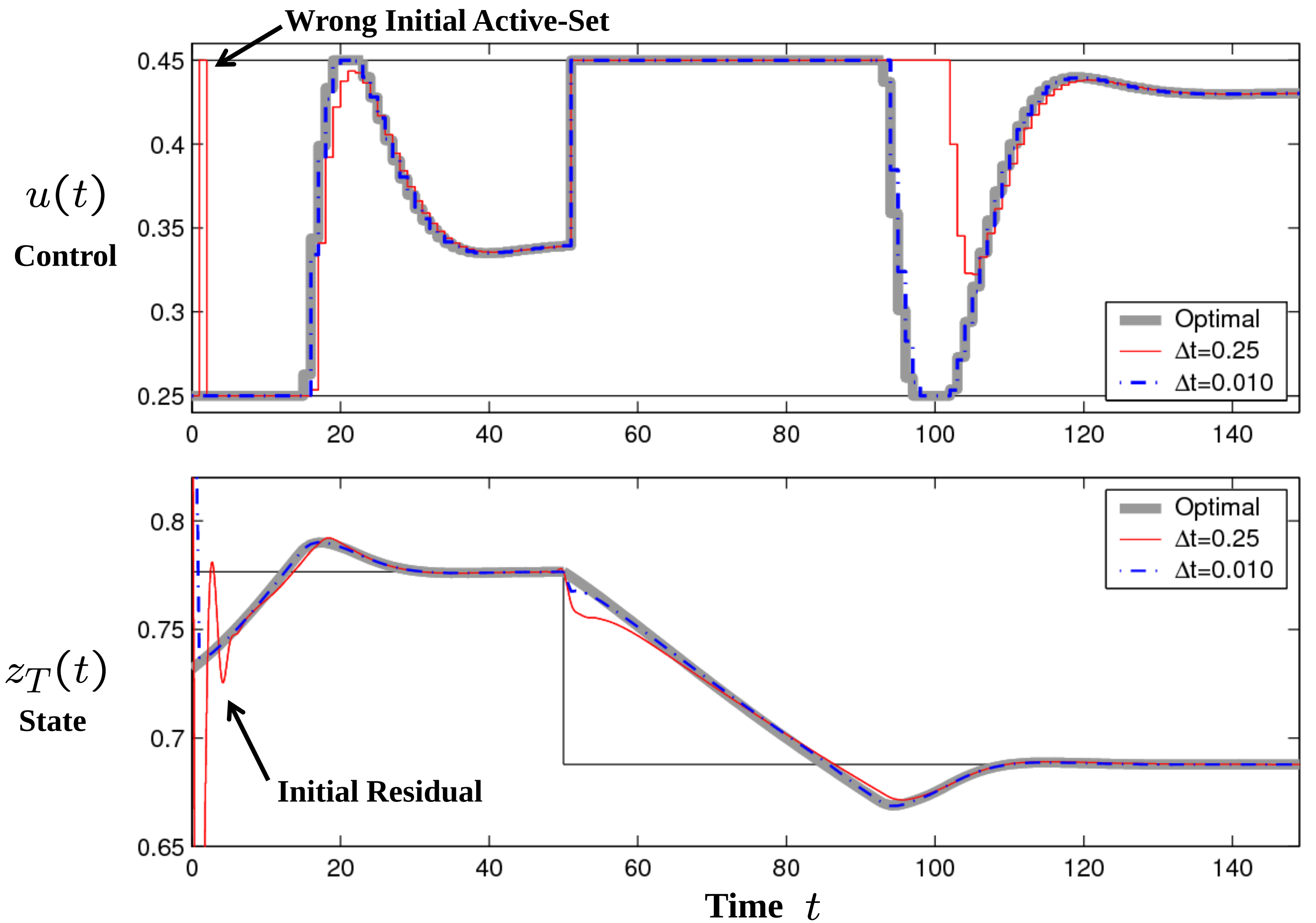
$N_L = 25, \ \rho = 100$



Sampling Time Restricted PGS Iterations, Warm-Smart Helps

Numerical Case Study

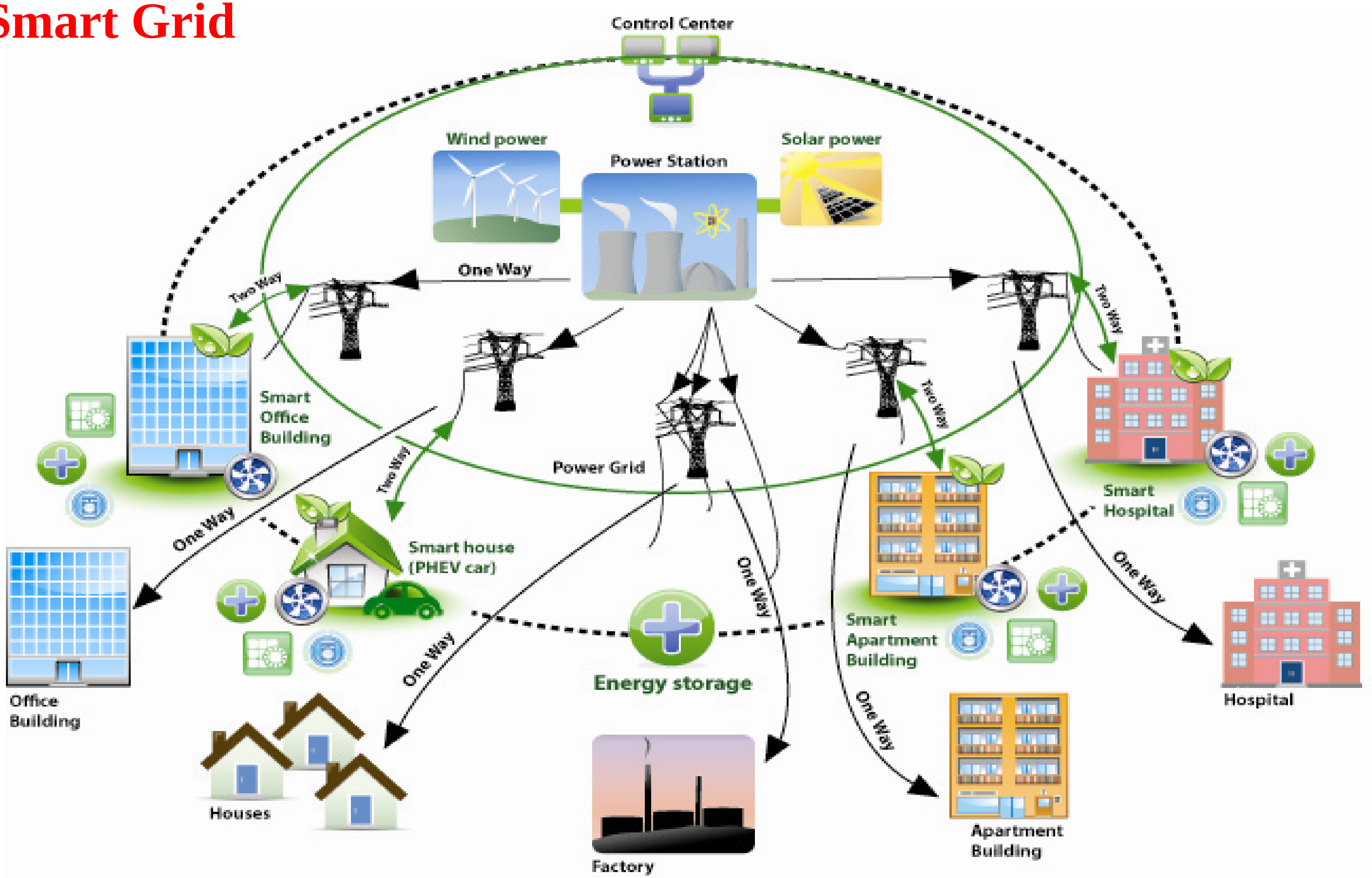
Optimal vs. Approximate Profiles



Future Work

Real-Time Grid Simulation

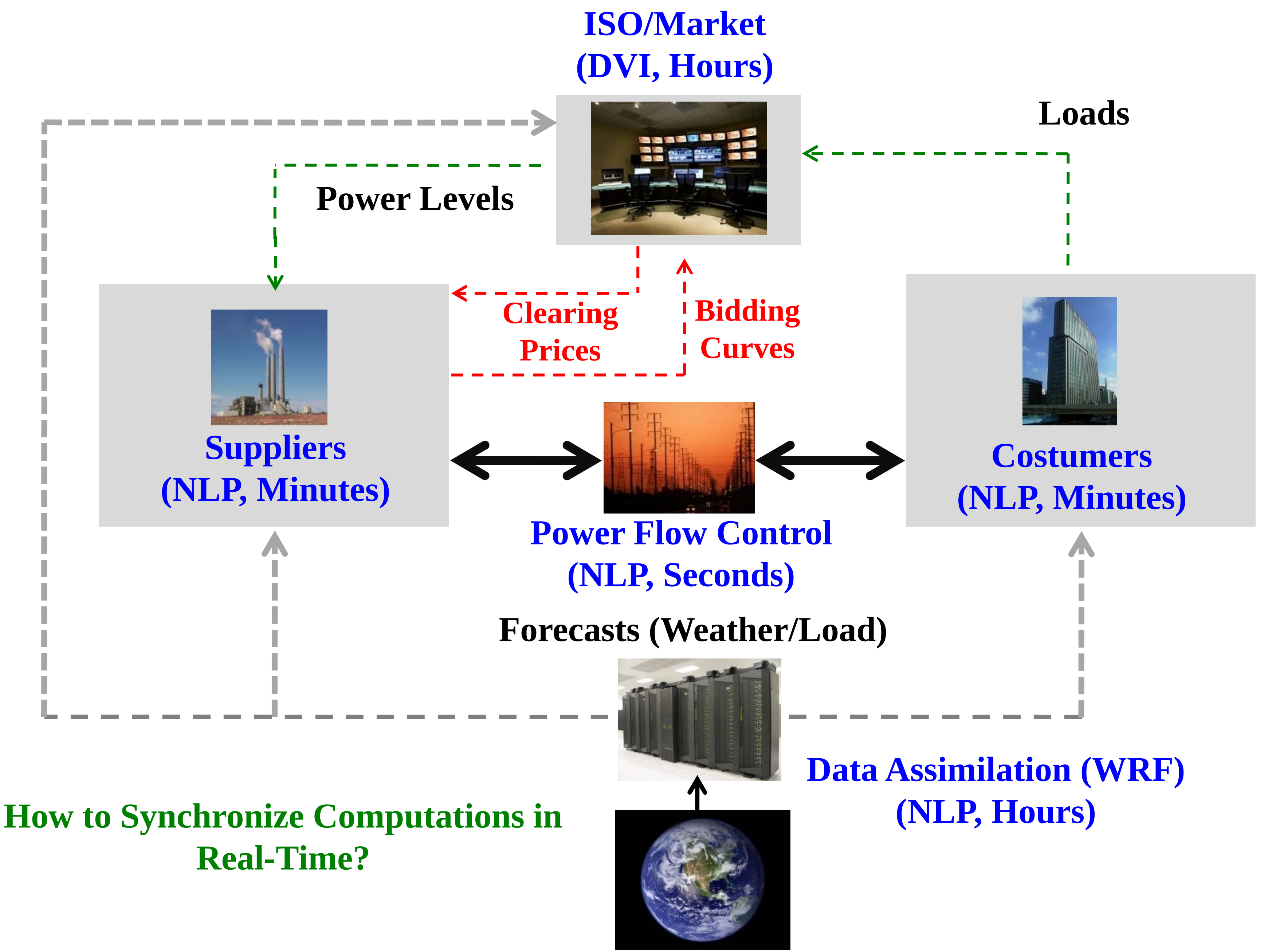
Smart Grid



Major Adoption of Renewable Resources (20-30%)

Deregulation, Highly <u>Distributed</u> Generation and Demands

Real-Time Grid Simulation



Conclusions and Future Work

Real-Time Optimization
- Interplays System Dynamics, Solution Accuracy, and Performance
- Manifold Tracking Algorithms

Generalized Equations
- Error Analysis
- Augmented Lagrangian and Projection Methods – Detection of Activity

Future Work
- Grid Simulator

  Other Strategies – Time-Stepping, Projected Gradients

  Distributed, Asynchronous, Warm-Starts, Exploit Periodicity

Real-Time Strategies for Control Architectures using Differential Variational Inequalities

Victor M. Zavala and Mihai Anitescu
Mathematics and Computer Science Division
Argonne National Laboratory
vzavala@mcs.anl.gov

Midwest Numerical Analysis Conference
April 24th, 2010