Multi-Scale Control of Energy Networks

Victor M. Zavala

Scalable Systems Laboratory
Department of Chemical & Biological Engineering
University of Wisconsin-Madison

Mathematics and Computer Science Division
Argonne National Laboratory

May, 2016
Multi-Scale Networks

Multiple Time Scales

Electrical Load

Ambient Temperature

U.S. Natural Gas Storage

You are Here
- Extremely low temperatures pushed most natural gas production and storage to urban areas to fulfill heating loads.

- California’s electric grid operator asked power generators to reduce their gas usage. Southern California Gas asked its customers to power down where possible.

- In Texas, which produces more natural gas than any other state, the Electric Reliability Council of Texas (ERCOT) called a state of emergency on Thursday out of concern there wouldn’t be enough gas for power plants. Earlier in the week ERCOT had asked to fire up two big coal-fired plants that are normally but on standby during the winter.

- In New England, tight supplies during the dreaded Polar Vortex caused the price of natural gas to soar 20-fold to more than $100 per thousand cubic feet. Gas was in such short supply in late January that New England’s grid operators told power generators to fire up not just their coal burners, but even their peaking plants that run on oil and even jet fuel.

- Total lost power generation capacity of 35GW, economic losses of 175 Million USD/hr
Multiple Spatial Scales

Time [hr]
- 0
- 5
- 10
- 15
- 20

Load [MW]
- 4
- 6
- 8
- 10
- 12
- 14
- 16

True

High Level

Mid Level

Low Level

Total Annual Energy Consumption kWh/m² block area
- 0 - 50
- 50 - 75
- 75 - 95
- 95 - 120
- 120 - 150
- 150 - 200
- 200 - 255
- 255 - 445
- 445 - 750
- 750 - 1,345
- 1,345 - 2,150
- > 2,150
Coupled Networks

\[ R_i = \underbrace{\pi_{n(i)}g_i}_{\text{Commitment}} + \underbrace{(G_i - g_i)\Pi_{n(i)}}_{\text{Real-Time}} \]

Natural gas demand set-points dictated by ISOs
Coupled Networks

Transport Equations for link $\ell \in \mathcal{L} := \mathcal{L}_p \cup \mathcal{L}_a$

\[
\frac{\partial p_\ell}{\partial t} + \frac{1}{A_\ell} \frac{\partial f_\ell}{\partial x} = 0
\]

\[
\frac{1}{A_\ell} \frac{\partial f_\ell}{\partial t} + \frac{\partial p_\ell}{\partial x} + \frac{8\lambda_\ell}{\pi^2 D_\ell^5} \frac{f_\ell|_{x=0}}{\rho_\ell} = 0
\]

\[
f_\ell|_{x=0} = f_\ell^{in}
\]

\[
f_\ell|_{x=L_\ell} = f_\ell^{out}
\]

\[
p_\ell|_{x=L_\ell} = \theta_{rec(\ell)}
\]

\[
p_\ell|_{x=0} = \theta_{snd(\ell)}, \quad \ell \in \mathcal{L}_p
\]

\[
p_\ell|_{x=0} = \theta_{snd(\ell)} + \Delta \theta_\ell, \quad \ell \in \mathcal{L}_a
\]

Conservation at node $n \in \mathcal{N}$

\[
\sum_{\ell: rec(\ell) = n} f_\ell^{out} + \sum_{i: sup(i) = n} s_i - \sum_{\ell: snd(\ell) = n} f_\ell^{in} - \sum_{j: dem(j) = n} d_j = 0
\]

Compression Power for link $\ell \in \mathcal{L}_A$

\[
P_\ell = f_\ell^{in} c_p T \left( \left( \frac{\theta_{snd(\ell)} + \Delta \theta_\ell}{\theta_{snd(\ell)}} \right)^\frac{\gamma-1}{\gamma} - 1 \right)
\]
• Under coordination **7% more gas is delivered**. At a gas price of 3$/MMBTU, the value of the additional gas delivered is **$1,070,000**.

• Under coordination the **revenue of the gas-fired power plants increases by 27%**, which corresponds to a total of **$800,000**.

**Key:** Under coordination, power plants behave as dispatchable gas loads.
Need to capture long-term transport/storage effects of gas infrastructure.
How do we *navigate* complexity?

How do we *coordinate* networks over multiple spatial and temporal scales?

What *information* should we *exchange* as well as where and when?

Can we design control architectures with *optimality/stability* guarantees?

How to we manage and leverage *multi-scale sensor* data?
Multi-Grid Schemes

Multi-Grid:
Computing paradigm for the solution of PDEs.
Only known paradigm that achieves perfect linear scaling.
Constructs hierarchy of fine to coarse system representations.
High-frequencies (e.g., local defects) “killed” by a smoother (i.e., Gauss-Seidel).
Low-frequencies (e.g., global defects) “killed” by a coarse operator.
Enables Management of Communication Latency and Asynchronous Tasks

Key Components of Multi-Grid:
Constructing convergent smoothers (distributed schemes).
Constructing operators that transfer information between levels.

Multi-Grid Provides Framework to Design Hierarchical Control Architectures
Structure of Optimal Control Problems

\[
\begin{align*}
\min_{u(\cdot)} & \quad \int_0^T \varphi(z(\tau), u(\tau)) d\tau \\
\text{s.t.} & \quad \dot{z}(\tau) = f(z(\tau), u(\tau)) \\
& \quad z(0) = \bar{z}
\end{align*}
\]

\[\downarrow \text{Lift Horizon}\]

\[
\begin{align*}
\min_{u_k(\cdot)} & \quad \sum_{k \in \mathcal{N}} \int_0^h \varphi(z_k(\tau), u_k(\tau)) d\tau \\
\text{s.t.} & \quad \dot{z}_k(\tau) = f(z_k(\tau), u_k(\tau)), \quad k \in \mathcal{N} \\
& \quad (\lambda_{k+1}) \quad z_{k+1}(0) = z_k(h), \quad k \in \mathcal{N}^- \\
& \quad z_0(0) = \bar{z}
\end{align*}
\]

\[\downarrow \text{Direct Transcription}\]

\[
\begin{align*}
\min_{u_{k,j}} & \quad \sum_{k \in \mathcal{N}} \sum_{j \in \mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}) \\
\text{s.t.} & \quad (\nu_{k,j+1}) \quad z_{k,j+1} = z_{k,j} + \delta f(z_{k,j+1}, u_{k,j+1}), \quad k \in \mathcal{N}, \quad j \in \mathcal{M} \\
& \quad (\lambda_k) \quad z_{k,0} = z_{k-1,m}, \quad k \in \mathcal{N}
\end{align*}
\]

\[\mathcal{M} := \{0..m-1\}, \quad z_{-1,m} := \bar{z}\]
Structure of Optimal Control Problems

\[ \min_{u_{k,j}} \sum_{k \in \mathcal{N}} \sum_{j \in \mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}) \]

s.t.
\[ (\nu_{k,j+1}) \ z_{k,j+1} = z_{k,j} + \delta f(z_{k,j+1}, u_{k,j+1}), \ k \in \mathcal{N}, j \in \mathcal{M} \]
\[ (\lambda_k) \ z_{k,0} = z_{k-1,m}, \ k \in \mathcal{N} \]

Compact Form
\[ \min_{u_k} \sum_{k \in \mathcal{N}} \phi(z_k, u_k) \]

s.t.
\[ (\nu_k) \ 0 = \chi(z_k, u_k), \ k \in \mathcal{N} \]
\[ (\lambda_k) \ \Pi_k z_k = \Pi_k z_{k-1}, \ k \in \mathcal{N} \]

where
\[ \phi(z_k, u_k) := \sum_{j \in \mathcal{M}} \varphi(z_{k,j+1}, u_{k,j+1}) \]
\[ \chi(z_k, u_k) := \begin{bmatrix} z_{k,1} - z_{k,0} - \delta f(z_{k,1}, u_{k,1}) \\ \vdots \\ z_{k,m} - z_{k,m-1} - \delta f(z_{k,m}, u_{k,m}) \end{bmatrix} \]
\[ \Pi_k z_k = z_{k,0} \quad \Pi_k z_{k-1} = z_{k-1,m} \]

KKT System
\[ 0 = \nabla_z \phi_k - \nabla_z \chi_k^T \nu_k - \Pi_k^T \lambda_k + \Pi_{k+1}^T \lambda_{k+1}, \ k \in \mathcal{N} \]
\[ 0 = \nabla_z \phi_{n-1} - \nabla_z \chi_{n-1}^T \nu_{n-1} - \Pi_{n-1}^T \lambda_{n-1} \]
\[ 0 = \nabla_u \phi_k - \nabla_u \chi_k^T \nu_k, \ k \in \mathcal{N} \]
\[ 0 = \chi(z_k, u_k), \ k \in \mathcal{N} \]
\[ 0 = \Pi_k z_k - \Pi_k z_{k-1}, \ k \in \mathcal{N} \]
Recursive Solution of KKT System

Stage KKT

\[
\begin{align*}
0 &= \nabla_z \phi_k - \nabla_z \chi_k^T \nu_k - \Pi_k^T \lambda_k + \Pi_{k+1}^T \lambda_{k+1} \\
0 &= \nabla_u \phi_k - \nabla_u \chi_k^T \nu_k \\
0 &= \chi(z_k, u_k) \\
0 &= \Pi_k z_k - \Pi_k z_{k-1}^\ell
\end{align*}
\]

Recall that \( \Pi_{k+1} z_k = z_{k,m} \)

Stage OCP

\[
\begin{align*}
\min_{u_k} \quad & \phi(z_k, u_k) + \lambda_{k+1}^\ell \Pi_{k+1} z_k \\
\text{s.t.} \quad & \nu_k = \chi(z_k, u_k) \\
& \lambda_k = \Pi_k z_k = \Pi_k z_{k-1}^\ell
\end{align*}
\]

Routine:

\[
(z_{k,m}^\ell+1, \lambda_k^\ell+1) \leftarrow \mathcal{P}_k(z_{k-1,m}^\ell, \lambda_{k+1}^\ell)
\]

Gauss-Seidel Scheme

I) GIVEN \( \bar{z} \), set \( \ell \leftarrow 0 \) and \( \lambda_k^\ell \leftarrow 0 \) for \( k = 0, \ldots, n - 1 \). FOR \( \ell = 0, \ldots, n_{GS} \):

II) FOR \( k = 0 \) SOLVE

\[
(z_{k,m}^{\ell+1}, \lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(\bar{z}, \lambda_{k+1}^\ell).
\]

III) FOR \( k = 1, \ldots, n - 2 \) SOLVE

\[
(z_{k,m}^{\ell+1}, \lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k-1,m}^{\ell+1}, \lambda_{k+1}^\ell).
\]

IV) FOR \( k = n - 1 \) SOLVE

\[
(z_{n-1,m}^{\ell+1}, \lambda_{n-1}^{\ell+1}) \leftarrow \mathcal{P}_{n-1}(z_{n-2,m}^{\ell+1}, 0).
\]

V) SET \( \ell \leftarrow \ell + 1 \) and RETURN TO Step II).
Recursive Solution of KKT System

\[
\begin{align*}
\min_{\mathbf{u}_k} & \quad \phi(z_k, u_k) + (\lambda_{k+1}^\ell)^T \Pi_{k+1} \mathbf{z}_k \\
\text{s.t.} & \quad (\lambda_k) \quad \Pi_k \mathbf{z}_k = \Pi_k \mathbf{z}_{k-1} 
\end{align*}
\]
**Connection with Receding-Horizon Control**

First Gauss-Seidel Sweep \((\lambda_k^\ell = 0)\)

\[
\begin{align*}
\max_{u_k} & \quad \phi(z_k, u_k) + (\lambda_k^\ell)\cdot T\cdot H_k z_k \\
\text{s.t.} & \quad \Pi_k z_k = \Pi_k^\ell z_k^{\ell - 1}
\end{align*}
\]

\[
\begin{align*}
\max_{u_{k+1}} & \quad \phi(z_{k+1}, u_{k+1}) + (\lambda_{k+1}^\ell)\cdot T\cdot H_{k+1} z_{k+1} \\
\text{s.t.} & \quad \Pi_{k+1} z_{k+1} = \Pi_{k+1}^\ell z_k^{\ell + 1}
\end{align*}
\]

**Implications:**
First Gauss-Seidel Sweep is Receding-Horizon (RH) Policy
If \(\lambda_k\) is Optimal and Unique \(\implies\) RH Policy is Optimal
Problem Boils Down to Estimating Stage Transition Duals \(\lambda_k\)
Coarsening (Two-Level)

**Issue:** Distributed schemes slow to converge.

**Key:** Accelerate using adjoint information from tractable coarse problem.

---

**Corrected Gauss-Seidel Scheme**

I) GIVEN $\bar{z}$, set counter $\ell \leftarrow 0$, set $z_{-1,m}^{\ell+1} \leftarrow \bar{z}$. FOR $\ell = 0, ..., n_{GS}$ DO:

II) SOLVE $(\lambda_0^\ell, ..., \lambda_{n-1}^\ell) \leftarrow \mathcal{P}_c(\bar{z})$ **Coarse Problem**

III) FOR $k = 0, ..., n - 2$ SOLVE

$(z_{k,m}^{\ell+1}, \lambda_k^{\ell+1}) \leftarrow \mathcal{P}_k(z_{k-1,m}^{\ell+1}, \lambda_{k+1}^\ell)$.

IV) FOR $k = n - 1$ SOLVE

$(z_{n-1,m}^{\ell+1}, \lambda_{n-1}^{\ell+1}) \leftarrow \mathcal{P}_{n-1}(z_{n-2,m}^{\ell}, 0)$.

V) SET $\ell \leftarrow \ell + 1$ and RETURN TO Step III).
Coarsening (Multi-Level)

\[ P_c \]

Coarse Level

\[ P_k(m/2) \]

Mid Level

\[ P_k(m) \]

Fine Level

\[ \lambda^\ell_{k+1} \]

\[ \lambda^\ell_{k+1} \]

\[ m/4 \]

\[ m/2 \]

\[ m \]

Time [hr]

Load [MW]
Objective: Find Cost-Optimal Generation/Storage Policies to Match Time-Varying Load

Constraints: Physical Bounds on Generation, Storage, and Flow
Multi-Scale Control (2-Level)

Net Load Frequencies

Load [MW] vs. Time [hr]

Load Error [MW] vs. Time [hr]

True Coarse
Multi-Scale Control (2-Level)

![Graphs showing multi-scale control](image)
Killing Low & High Frequencies
Adjoint Correction

![Graphs showing adjoint correction over time for GS (It 1), GS (It 2), GS (It 3), and GS (It 4).]
Multigrid Control (2-Level vs. 3-Level)

Load Coarsening \((m_c = m/4 = 16)\)

Load Coarsening \((m_c = m/16 = 4)\)
Multigrid Control (2-Level vs. 3-Level)

**Adjoints** ($m/4$ vs. $m/16$)

**Adjoint Error** (2-Level vs. 3-Level)

Multigrid Enables Systematic Transition Between Scales
Alternative Recursions

Long-Horizon Problem

$$\min_{u_k} \sum_{k \in \mathcal{N}} \phi(z_k, u_k)$$

s.t. $$(\lambda_k) \quad \Pi_k z_k = \Pi_k z_{k-1}, \; k \in \mathcal{N}$$

Lagrangian Dual Decomposition

$$\mathcal{L}(z_0..z_{n-1}, \lambda_0..\lambda_{n-1}) := \sum_{k \in \mathcal{N}} \phi(z_k, u_k) - \sum_{k \in \mathcal{N}} \lambda_k^T (\Pi_k z_k - \Pi_k z_{k-1})$$

$$\max_{\lambda_k} \min_{z_k} \mathcal{L}(z_0..z_{n-1}, \lambda_0..\lambda_{n-1})$$

ADMM

$$\mathcal{L}_A(z_0, z_1, ..., z_{n-1}, \lambda_0...\lambda_{n-1}) := \mathcal{L}(\cdot) + \frac{\rho}{2} \sum_{k \in \mathcal{N}} \|\Pi_k z_k - \Pi_k z_{k-1}\|^2$$

$$z_0^{\ell+1} = \arg\min_z \mathcal{L}_A(z, z_1^{\ell}, ..., z_{n-1}^{\ell}, \lambda_0^{\ell}...\lambda_{n-1}^{\ell})$$

$$z_1^{\ell+1} = \arg\min_z \mathcal{L}_A(z_0^{\ell+1}, z, ..., z_{n-1}^{\ell}, \lambda_0^{\ell}...\lambda_{n-1}^{\ell})$$

$$\vdots$$

$$z_{n-1}^{\ell+1} = \arg\min_z \mathcal{L}_A(z_0^{\ell+1}, z_1^{\ell+1}, ..., z, \lambda_0^{\ell}...\lambda_{n-1}^{\ell})$$

$$\lambda_k^{\ell+1} = \lambda_k^\ell + \rho \left( \Pi_k z_k^{\ell+1} - \Pi_k z_k^{\ell+1} \right), \; k \in \mathcal{N}$$

ADMM can be implemented using Gauss-Seidel and Jacobi schemes.
ADMM communicates adjoint & state information but eliminates *duality gap*. 
Multi-Scale State Estimation (Data Assimilation)

Want an estimator that uses sensor information emanating at multiple scales. Accumulating high- and low-frequency data is intractable. How to Compress/Summarize?

Structure of MHE problem

\[
\min_{z(\cdot), u(\cdot)} \frac{1}{2} \mu (z(0) - \bar{z})^T (z(0) - \bar{z}) \\
+ \int_0^T \varphi(z(\tau), u(\tau)) d\tau \\
\text{s.t. } \dot{z}(\tau) = f(z(\tau), u(\tau)), \ \tau \in [0, T].
\]

\[
\min_{z_k, u_k} \frac{1}{2} \mu (\Pi_0 z_0 - \Pi_0 z_{-1})^T (\Pi_0 z_0 - \Pi_0 z_{-1}) \\
+ \sum_{k \in \mathcal{N}} \phi(z_k, u_k) \\
\text{s.t. } (\lambda_k) \ \Pi_k z_k = \Pi_k z_{k-1}, \ k \in \mathcal{N} \setminus \{0\}.
\]
Moving Horizon Estimation

\[
\min_{z_k, u_k} \frac{1}{2} \mu (\Pi_k z_k - \Pi_k z_{k-1})^T (\Pi_k z_k - \Pi_k z_{k-1}) + \phi(z_k, u_k) \quad \mathcal{P}_k
\]

\[
\min_{z_{k+1}, u_{k+1}} \frac{1}{2} \mu (\Pi_{k+1} z_{k+1} - \Pi_{k+1} z_k)^T (\Pi_{k+1} z_{k+1} - \Pi_{k+1} z_k) + \phi(z_{k+1}, u_{k+1}) \quad \mathcal{P}_{k+1}
\]

\[
\min_{z_{n}, u_{n}} \frac{1}{2} \mu (\Pi_n z_n - \Pi_n z_{n-1})^T (\Pi_n z_n - \Pi_n z_{n-1}) + \phi(z_n, u_n) \quad \mathcal{P}_n
\]

In MHE, information is propagated (summarized) in arrival cost.

\[
\min_{z_k, u_k} \phi(z_k, u_k) + (\lambda_{k+1})^T \Pi_{k+1} z_k
\]

\[
\text{s.t. } (\lambda_k) \quad \Pi_k z_k = \Pi_k z_{k-1}
\]

\[
\min_{z_{k+1}, u_{k+1}} \phi(z_{k+1}, u_{k+1}) + (\lambda_{k+2})^T \Pi_{k+2} z_{k+1}
\]

\[
\text{s.t. } (\lambda_{k+1}) \quad \Pi_{k+1} z_{k+1} = \Pi_{k+1} z_k
\]

\[
\min_{z_n, u_n} \phi(z_n, u_n)
\]

\[
\text{s.t. } (\lambda_n) \quad \Pi_n z_n = \Pi_n z_{n-1}
\]

In multi-scale MHE, information is propagated through initial state.
Moving Horizon Estimation
Spatial Schemes

\[
\min_{z_s} \sum_{s \in S} \varphi_s(z_s)
\]

s.t. \( \Pi_{s,s'} z_s = \Pi_{s,s'} z_{s'}, \ s \in S, s' \in S, (\lambda_{s,s'}) \)

\( \Pi_{s,s'}, \Pi_{s,s'} \) encode connectivity information.

Adjoint \( \lambda_{s,s'} \) transfer information between subnetworks/nodes \( s, s' \).

Can construct coarse network representations to estimate \( \lambda_{s,s'} \).

Observation: An infinite number of Gauss-Seidel (GS) schemes can be devised.
Global effects propagate slowly in Peer-to-Peer (GS) scheme.
Coarsening captures global effects, GS captures local effects.
From Nodes to Hubs?

Hub 1
Hub 2
Hub 3
Hub 4
Hub 5
Multi-Scale Control of Energy Networks

Victor M. Zavala

Scalable Systems Laboratory
Department of Chemical & Biological Engineering
University of Wisconsin-Madison

Mathematics and Computer Science Division
Argonne National Laboratory

May, 2016
Why on Earth Would Gauss-Seidel Work?

Consider **Two-Stage Problem**

\[
\begin{align*}
\min_{u_k, u_{k+1}} & \quad \phi(z_k, u_k) + \phi(z_{k+1}, u_{k+1}) \\
\text{s.t.} & \quad 0 = \chi(z_k, u_k) \\
& \quad \Pi_k z_k = \Pi_k z_{k-1} \\
& \quad 0 = \chi(z_{k+1}, u_{k+1}) \\
& \quad \Pi_{k+1} z_{k+1} = \Pi_{k+1} z_k (\lambda_{k+1})
\end{align*}
\]

Consider **Linearization** of Cost-To-Go

\[
\begin{align*}
\min_{u_k} & \quad \phi(z_k, u_k) + Q(\Pi_{k+1} z_k^\ell) + \frac{\partial Q(\Pi_{k+1} z_k^\ell)^T}{\lambda_{k+1}(\Pi_{k+1} z_k^\ell) = \lambda_{k+1}^\ell} (\Pi_{k+1} z_k - \Pi_{k+1} z_k^\ell) \\
\text{s.t.} & \quad 0 = \chi(z_k, u_k) \\
& \quad \Pi_k z_k = \Pi_k z_{k-1}
\end{align*}
\]

\[
\begin{align*}
\min_{u_k} & \quad \phi(z_k, u_k) + (\lambda_{k+1}^\ell)^T \Pi_{k+1} z_k \\
\text{s.t.} & \quad 0 = \chi(z_k, u_k) \\
& \quad \Pi_k z_k = \Pi_k z_{k-1}
\end{align*}
\]

**Cost-To-Go**

\[
Q(\Pi_{k+1} z_k^\ell) := \min_{u_{k+1}} \phi(z_{k+1}, u_{k+1})
\]

\[
\begin{align*}
\Pi_{k+1} z_{k+1} = \Pi_{k+1} z_k (\lambda_{k+1})
\end{align*}
\]

**Implications:**

Gauss-Seidel is a Gradient Scheme for Composite Functions

Can Estimate \( \partial Q \) (i.e., \( \lambda_k \)) Using DP or Coarsening

Adjoint acts as price of future information