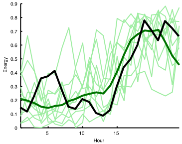
Computational Methods for Distributed Controller Design in the Smart Grid

J. Zico Kolter
School of Computer Science
Carnegie Mellon University

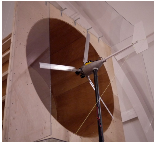
Joint work with Matt Wytock

February 5, 2014

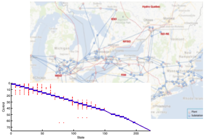
Recent and ongoing projects



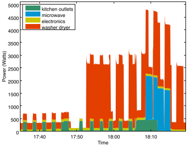
Probabilistic power forecasting



Learning micro wind turbine control

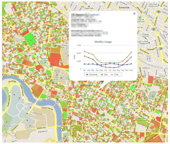


Fast algorithms for decentralized wide area power control



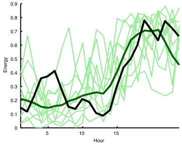
Energy disaggregation

Residential energy data collection

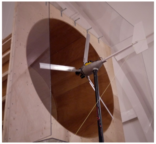


City-level energy modeling and visualization

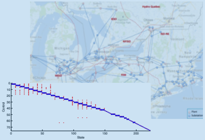
Recent and ongoing projects



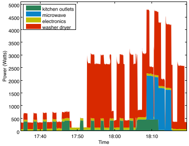
Probabilistic power forecasting



Learning micro wind turbine control

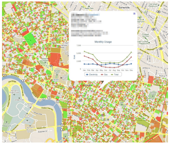


Fast algorithms for decentralized wide area power control



Energy disaggregation

Residential energy data collection



City-level energy modeling and visualization

From big data to big control

Big Data

Many data points
Many features
High data velocity

From big data to big control

Big Data

Many data points
Many features
High data velocity

Big Control

Big data +
Many decisions
Fast dynamics
Communication limits

From big data to big control

Big Data

Many data points
Many features
High data velocity

Sparsity:
Simpler models
Learn relevant structure
Faster computation

Big Control

Big data +
Many decisions
Fast dynamics
Communication limits

From big data to big control

Big Data

Many data points
Many features
High data velocity

Sparsity:
Simpler models
Learn relevant structure
Faster computation

Big Control

Big data +
Many decisions
Fast dynamics
Communication limits

Sparsity:
Simpler communication
Fixed structure
Computation?

From big data to big control

Big Data

Many data points
Many features
High data velocity

Sparsity:
Simpler models
Learn relevant structure
Faster computation

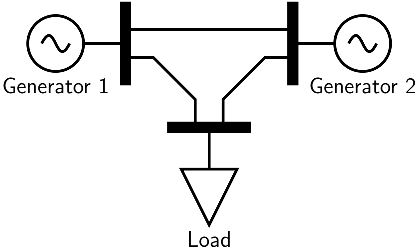
Big Control

Big data +
Many decisions
Fast dynamics
Communication limits

Sparsity:
Simpler communication
Fixed structure
Computation?

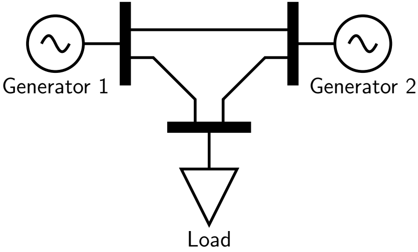
This work: towards bringing CS big data approaches into the computational side of control

Wide area control of power systems



$$x = \begin{bmatrix} \theta_1 \\ \dot{\theta}_1 \\ \theta_2 \\ \dot{\theta}_2 \end{bmatrix}, \quad u = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

Wide area control of power systems



$$x = \begin{bmatrix} \theta_1 \\ \dot{\theta}_1 \\ \theta_2 \\ \dot{\theta}_2 \end{bmatrix}, \quad u = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

$$K = \begin{bmatrix} {\color{blue}\bullet} & {\color{blue}\bullet} & {\color{red}\bullet} & {\color{red}\bullet} \\ {\color{red}\bullet} & {\color{red}\bullet} & {\color{blue}\bullet} & {\color{blue}\bullet} \end{bmatrix}$$

${\color{blue}\bullet}$ = "local" control

${\color{red}\bullet}$ = "wide area" control

Decentralized control

- Huge area of research dating back to 60s: linear decentralized control known to be suboptimal (Witenhausen, 1962); survey of early work (Sandell Jr. et al., 1978)

- Non-convex for typical control objectives ($\mathcal{H}_2$, $\mathcal{H}_\infty$)

- A great deal of interest in recent years: special systems (e.g., quadratic invariance) where optimal decentralized control *is* sparse (Rotkowitz and Lall, 2006); convexifications using different control objectives (Dvijotham et al., 2013) or restricted Lyapunov functions (Schuler et al., 2013); non-convex optimization and ℓ_1 penalty (Lin et al., 2013)

- *But*, virtually no large-scale algorithms

Problem setup (following Lin et al., 2013)

- Continuous time linear dynamical system

$$\dot{x}(t) = Ax(t) + Bu(t) + W^{1/2}\epsilon(t)$$

- LQR/$\mathcal{H}_2$ objective for state feedback controller $u(t) = Kx(t)$

$$J(K) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \big(x(t)^T Q x(t) + u(t) R u(t)\big)\, dt$$

$$= \begin{cases} \operatorname{tr} PW & A + BK \text{ stable} \\ \infty & \text{otherwise} \end{cases}$$

$$(A+BK)^T P + P(A+BK) + Q + K^T R K = 0$$

- Optimization objective with ℓ_1 term

$$\underset{K}{\text{minimize}}\, J(K) + \lambda \|K\|_1$$

(a difficult optimization problem)

Computational challenges

- Taking inspiration from sparse big-data methods: Newton Coordinate Descent approaches (e.g. Hsieh et al., 2011, Tseng and Yun, 2009)

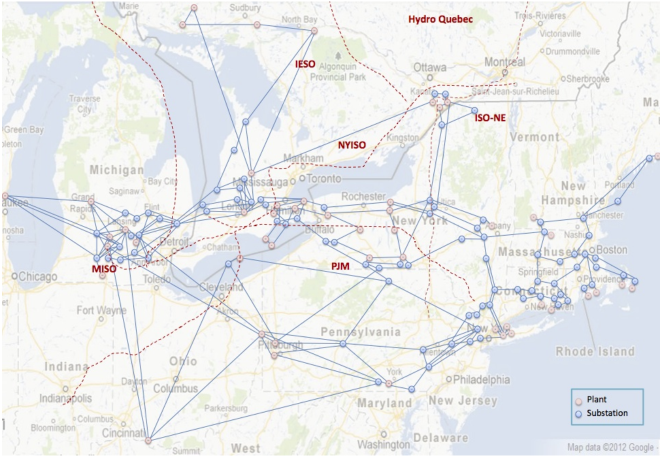
- Iteratively form quadratic approximation and minimize over search direction Δ

$$\mathrm{tr}(\nabla_K J(K))^T \Delta + \mathrm{vec}(\Delta)^T (\nabla_K^2 J(K)) \, \mathrm{vec}(\Delta) + \lambda \|K + \Delta\|_1$$

  where this step can exploit recent advances in fast ℓ_1 optimization

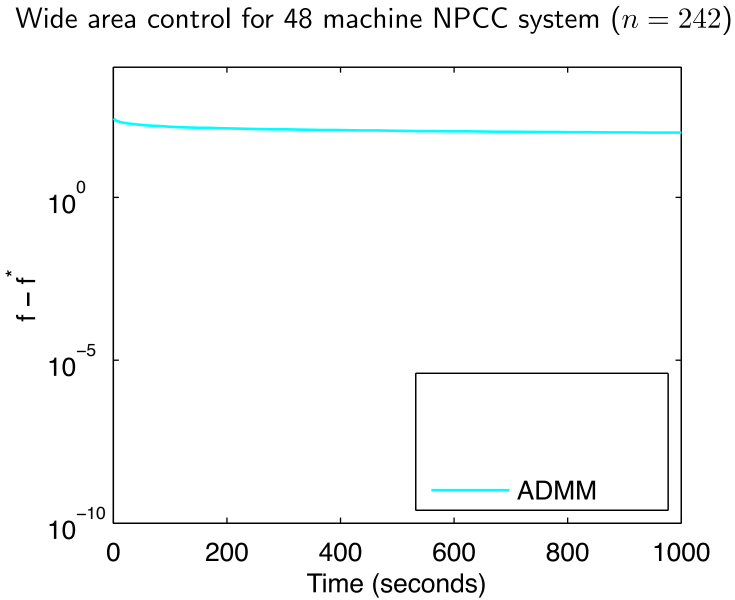
- But, Hessian products in particular are computationally intensive; majority of the work focuses on bringing this from $O(n^3) \rightarrow O(n)$

Power system wide area control

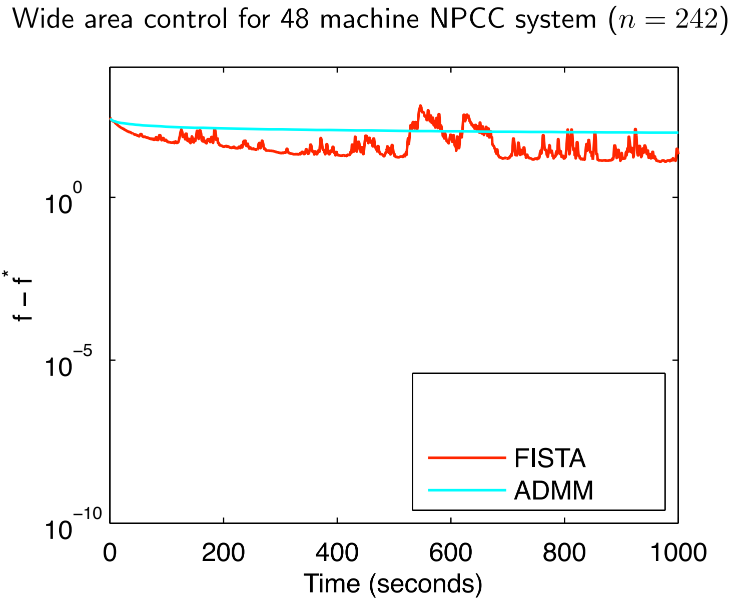


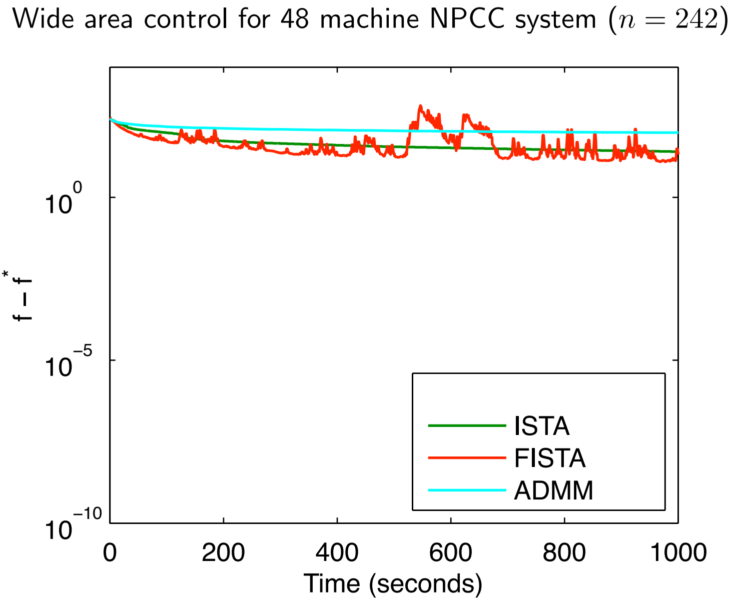
NPCC 48 machine, 140 bus system (Chow et al, 1995)

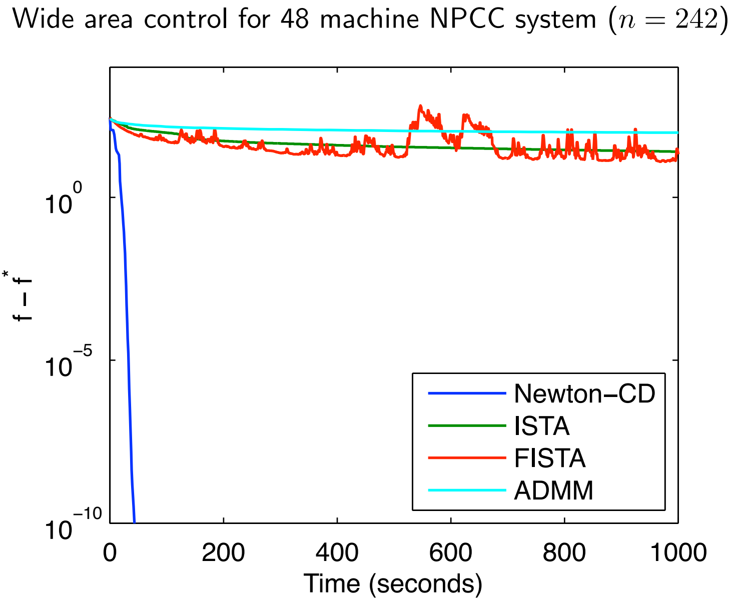
24 machines equipped with exciters and turbine governors, $n = 242$



Wide area control for 48 machine NPCC system ($n = 242$)
10^{0}
10^{-5}
10^{-10}
f − f*
0
200
400
600
800
1000
Time (seconds)
ADMM

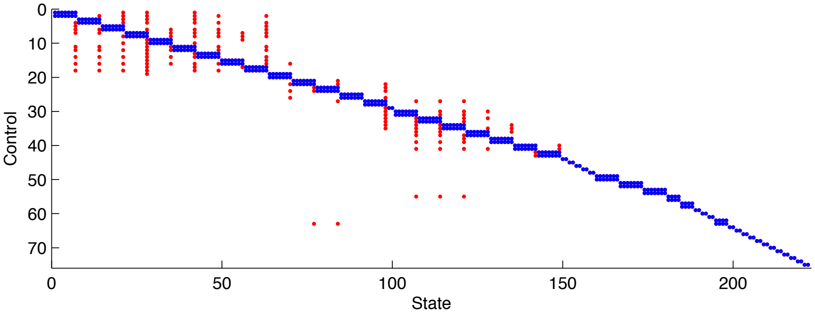

Wide area control for 48 machine NPCC system ($n = 242$)
f – f*
10^0
10^−5
10^−10
0
200
400
600
800
1000
Time (seconds)
FISTA
ADMM


Wide area control for 48 machine NPCC system ($n = 242$)
f − f*
10^{0}
10^{-5}
10^{-10}
0
200
400
600
800
1000
Time (seconds)
ISTA
FISTA
ADMM



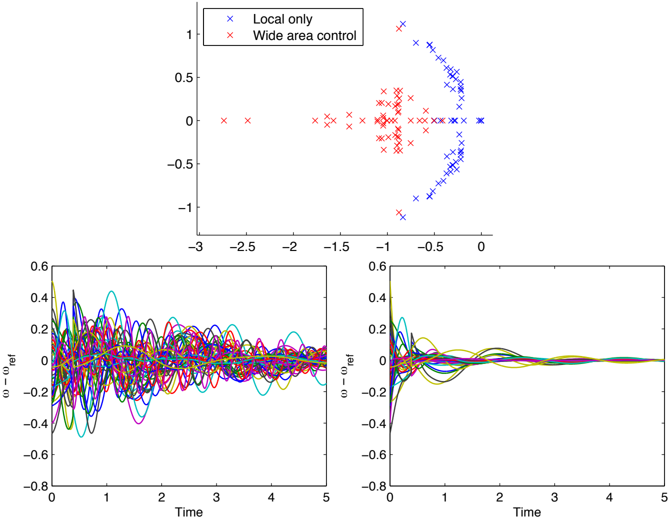
Wide area control for 48 machine NPCC system ($n = 242$)
$f - f^*$
10^0
10^{-5}
10^{-10}
0
200
400
600
800
1000
Time (seconds)
Newton−CD
ISTA
FISTA
ADMM

NPCC 48 machine system



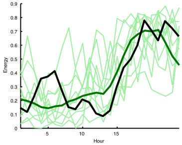
Plot of controller sparsity

NPCC 48 machine system

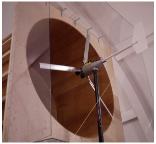


Eigenvalues and regulation of intial / wide area control systems

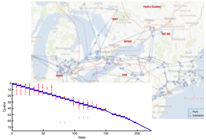
Recent and ongoing projects



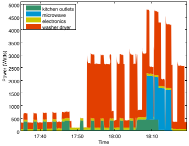
Probabilistic power forecasting



Learning micro wind turbine control

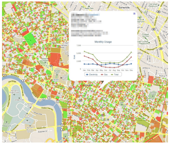


Fast algorithms for decentralized wide area power control



Energy disaggregation

Residential energy data collection



City-level energy modeling and visualization