REAL-TIME PRICE FORECAST WITH BIG DATA

A STATE SPACE APPROACH

Lang Tong (PI), Robert J. Thomas, Yuting Ji, and Jinsub Kim

School of Electrical and Computer Engineering, Cornell University

Jie Mei, Georgia Institute of Technology

August 7, 2013

CERTS Review

DATA QUALITY AND ITS EFFECTS ON MARKET OPERATIONS

DENY OF SERVICE ATTACK ON REAL-TIME ELECTRICITY MARKET
COVER UP PROTECTION AGAINST TOPOLOGY ATTACK

Lang Tong (PI), Robert J. Thomas, and Jinsub Kim

Cornell University

August 8, 2013

CERTS Review

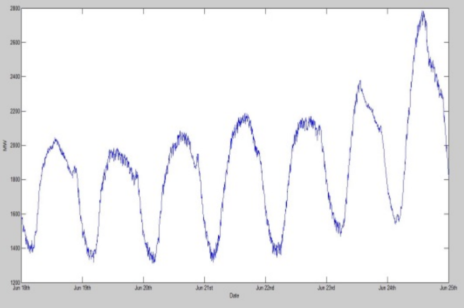
Project overview

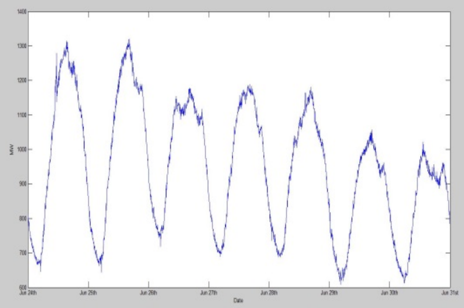
- Objectives
  - Accurate short-term probabilistic forecasting of real-time LMP.
  - Incorporate real-time measurements (e.g. SCADA/PMU).
  - Scalable computation techniques.
- Summary of results
  - A real-time LMP model with forecasting and measurement uncertainties.
  - A Markov chain abstraction of real-time LMP computation.
  - Monte Carlo sampling techniques.
  - Preliminary simulations.

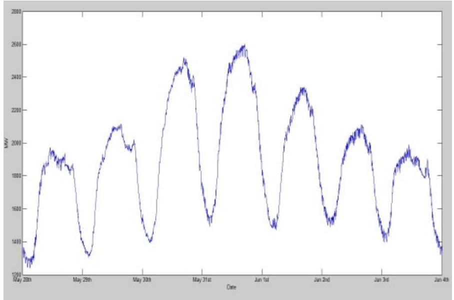
Outline

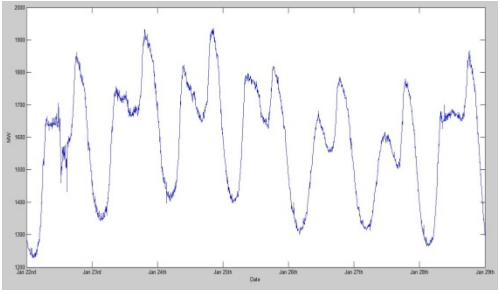
- Motivation
  - Load and real-time LMP as random processes
  - Benchmark techniques
  - Comments on the state of the art.
- Probabilistic forecasting of real-time LMP
- Simulation studies
- Summary and future work

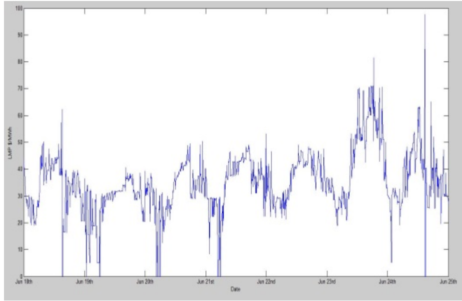
Sample paths of load and LMP

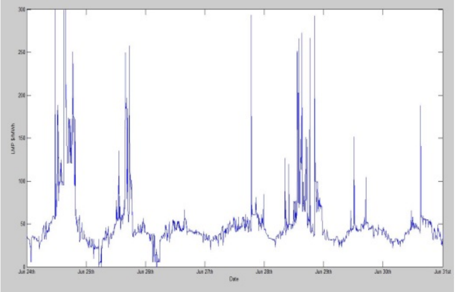


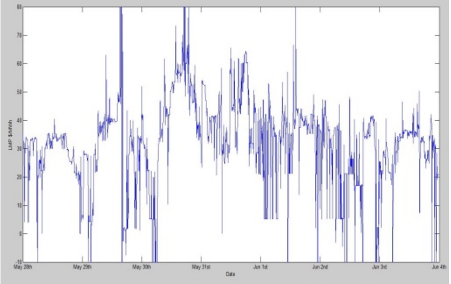


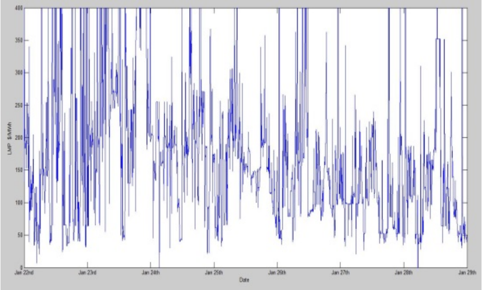




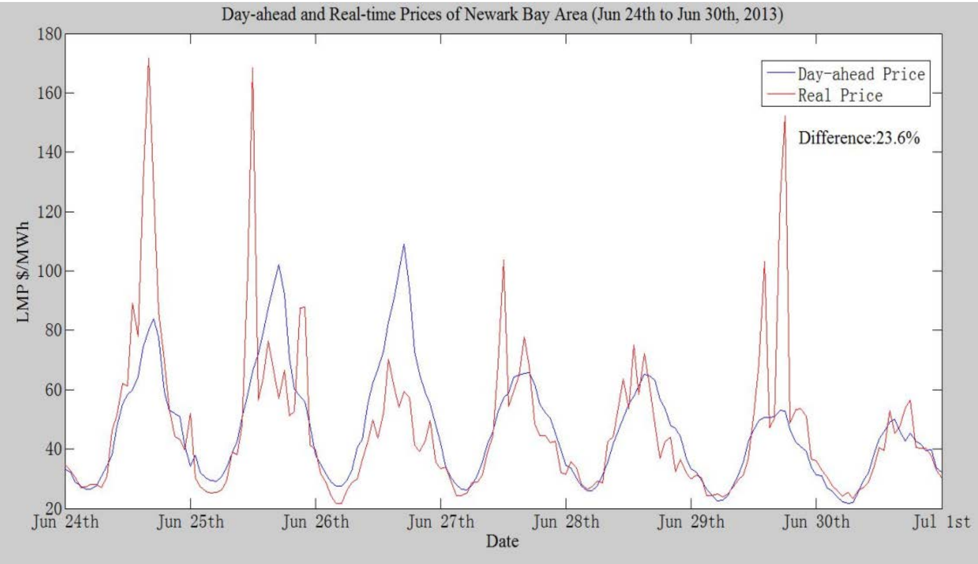


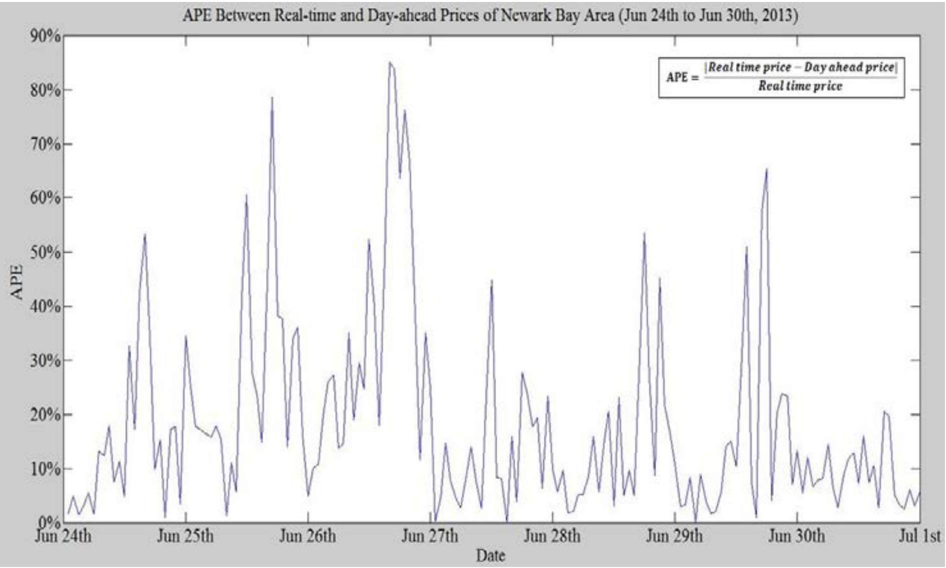






Day ahead vs. real-time

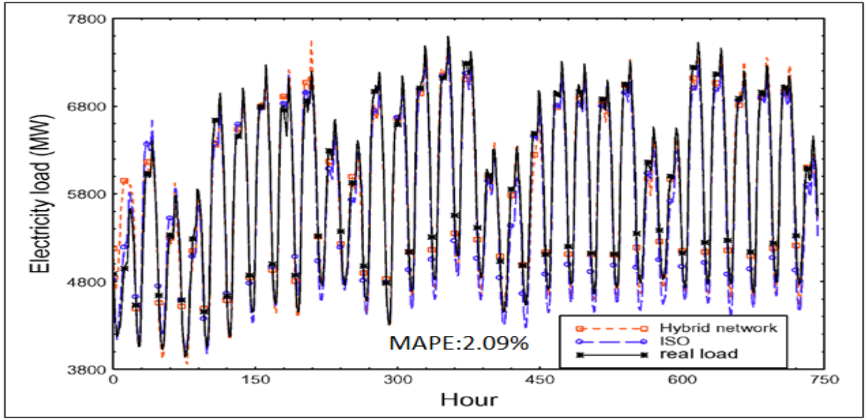


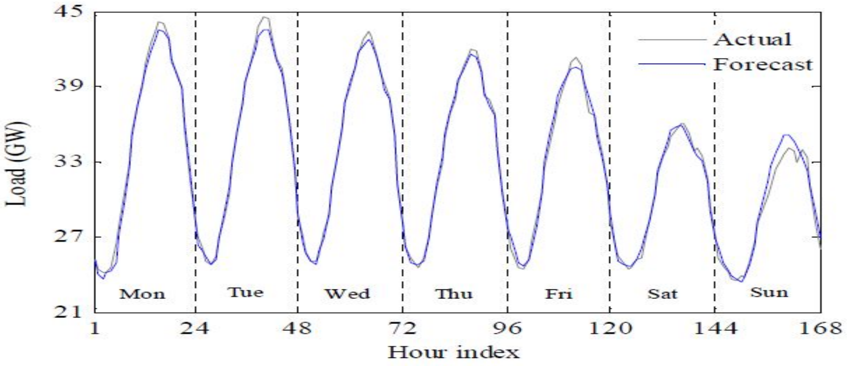


Benchmark techniques

- Time series
  - ARMA, ARIMA, ARMAX, GARCH
- Machine learning
  - Neural networks, support vector machines (SVM)
- Hybrid techniques
  - Jump (switching) models

Load forecasting

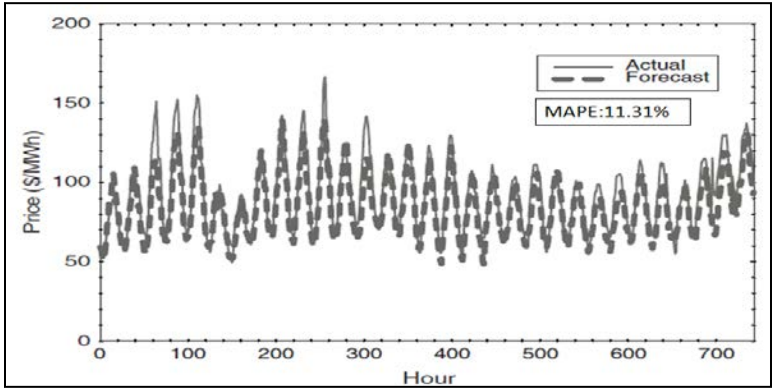


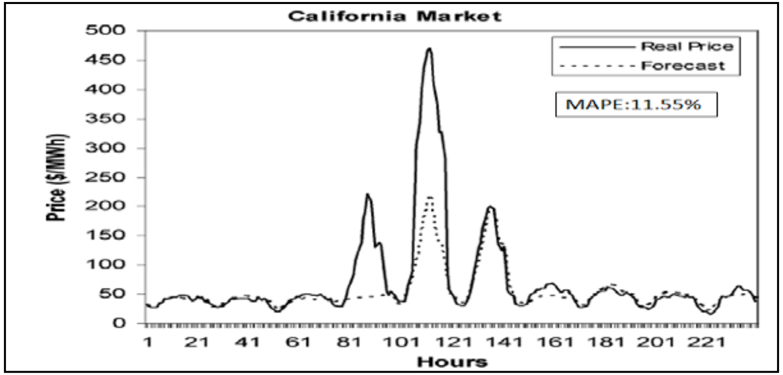


Forecasting Method	MAPE
ANN [1] ANN [2]	1.46% 1.24%
SVM [3]	2.09%

[1] P. Mandal, T. Senjyu, N. Urasaki, and T. Funabashi, "A neural network based several-hour-ahead electric load forecasting using similar days approach," Int. Journal. of Electric Power and Energy System, vol. 28, no. 6, pp. 367–373, July 2006.
[2] P. Mandal, T. Senjyu, N. Urasaki, and T. Funabashi, "Short-Term Price Forecasting for Competitive Electricity Market," Power Symposium, 2006. NAPS 2006. 38th North.
[3] S. Fan, C. Mao and L. Chen, "Next-day electricity-price forecasting using a hybrid network," IET Generation, Transmission & Distribution, Volume 1, Issue 1, January, 2007.

Day ahead LMP forecasting

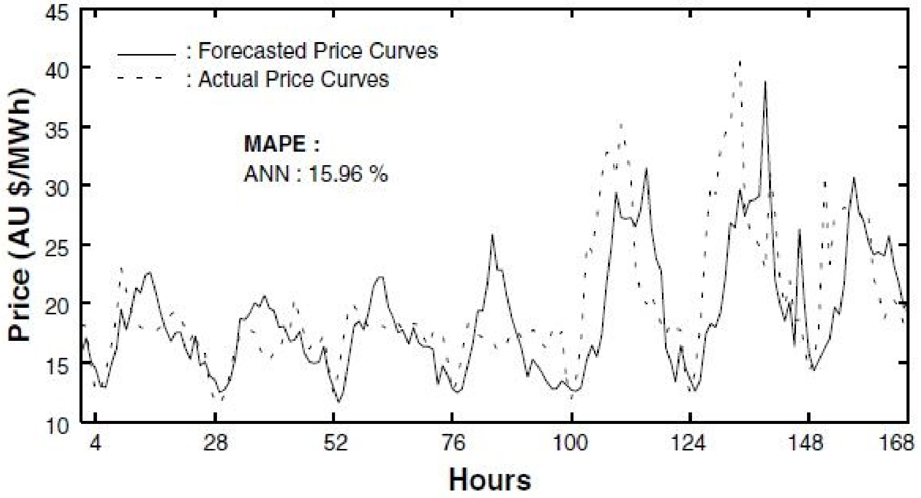




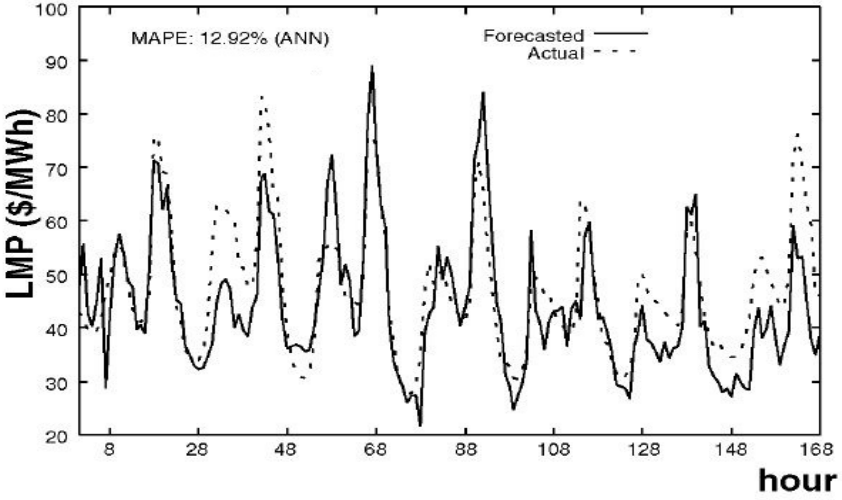
Forecasting Method	MAPE
ANN [1] ANN [2] ANN [4]	15.96% 12.92% 8.67%
SVM [3]	11.31%
GARCH [5]	11.55%
ARMAX [6]	9.01%

[4] P. Mandal, T. Senjyu, N. Urasaki, and T. Funabashi, “A neural network based several-hour-ahead electric load forecasting using similar days approach,” Int. Journal. of Electric Power and Energy System, vol. 28, no. 6, pp. 367–373, July 2006.
[5] R. Garcia, J. Contreras, “A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices,” Int. Journal of Electric Power and Energy System, vol. 20, no. 2, MAY 2005.
[6] J. Zhang, C. Yu, G. Hou, “Application of Chaotic Particle Swarm Optimization in the Short-term Electricity Price Forecasting,” Int. Journal of Electric Power and Energy System, vol. 28, no. 6, pp. 367–373, July 2006.

Real-time LMP forecasting



Actual and forecast six-hours-ahead electricity price for Victorian electricity market (Nov 1st to 7th, 2003).



Actual and forecast six-hours-ahead electricity price for PJM market (Jan 8th to 14th, 2006).

[4] F. Sarafraz, H. Ghasemi, H. Monsef, "Locational Marginal Price Forecasting by Locally
Linear Neuro-Fuzzy Model," EEEIC, 2011 10th, Rome

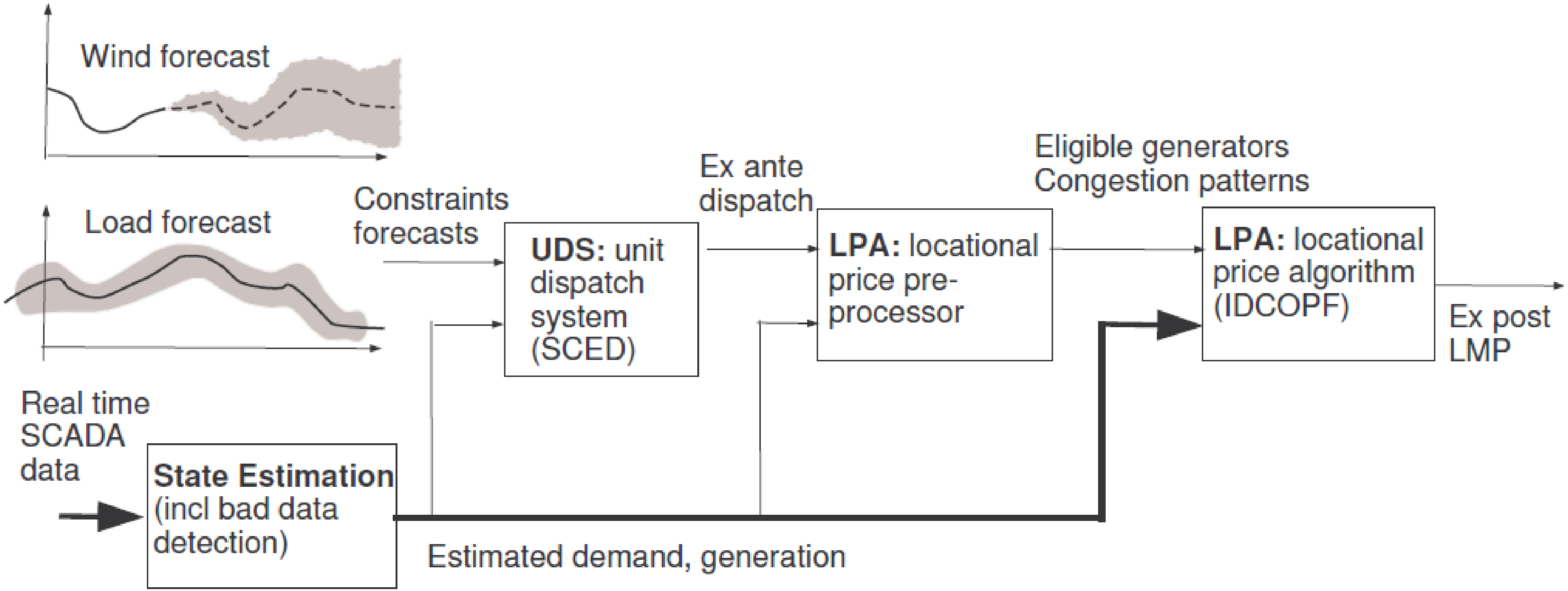
Summary of the state-of-the-art

- Extensive literature:
  - Mostly black box techniques
  - Primarily providing point forecasting
  - Rarely deal with on LMP and network effects
- Extremely accurate load forecasting (1-3%)
- Relatively poor price forecasting (10-20%)

Outline

- Motivation
- Probabilistic forecasting of real-time LMP
  - A stylized real-time ex post LMP model
  - Information structure
  - LMP states and Markov chain representations
  - Probabilistic forecast
- Preliminary simulations
- Summary and future work

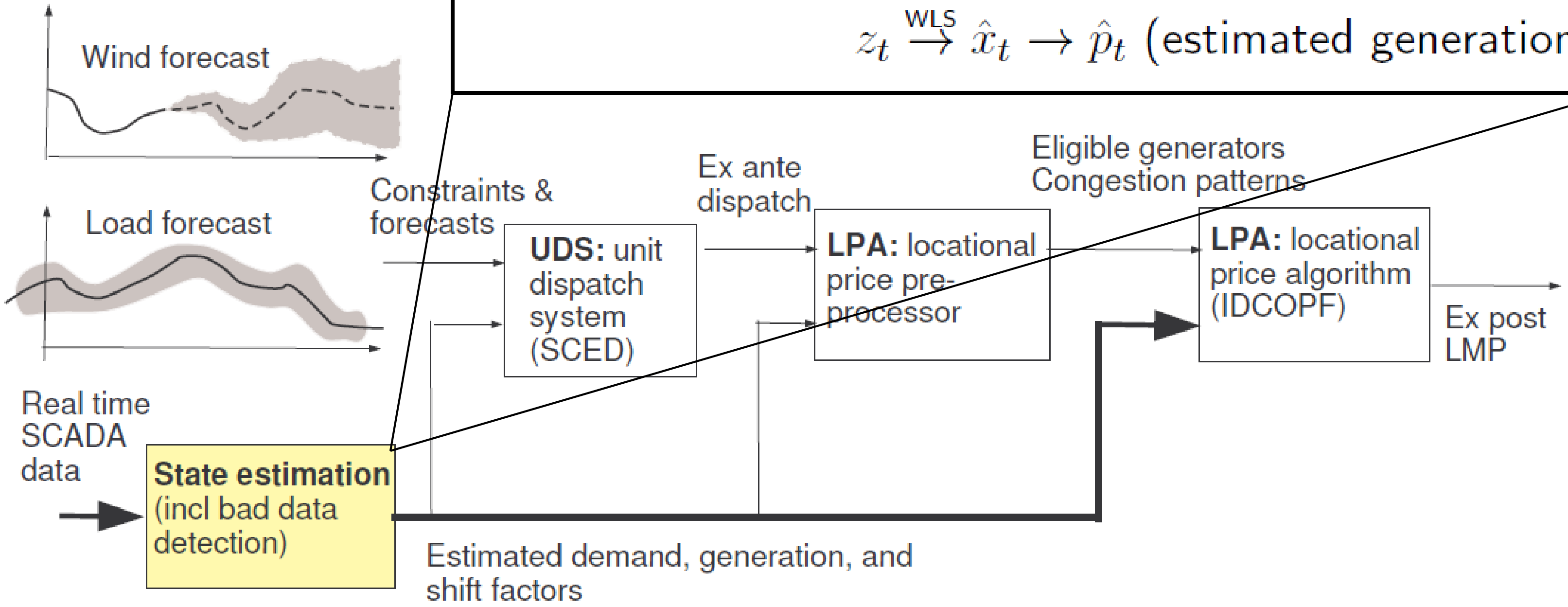
A stylized real-time ex post LMP model



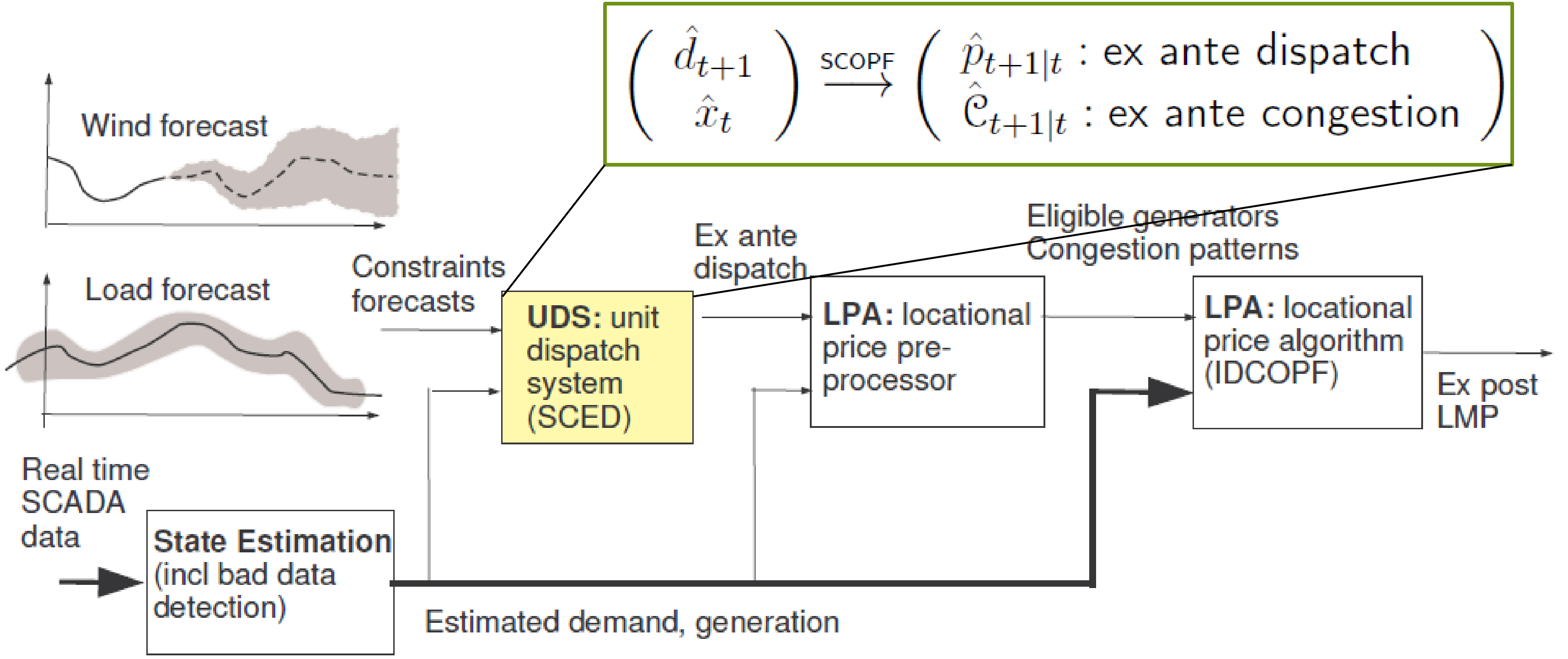
State estimation

- Power system as a graph: $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$
- SCADA measurement model: $z_t = h(x_t, \mathcal{G}) + e_t$

$$z_t \overset{\text{WLS}}{\longrightarrow} \hat{x}_t \rightarrow \hat{p}_t \text{ (estimated generation)}$$

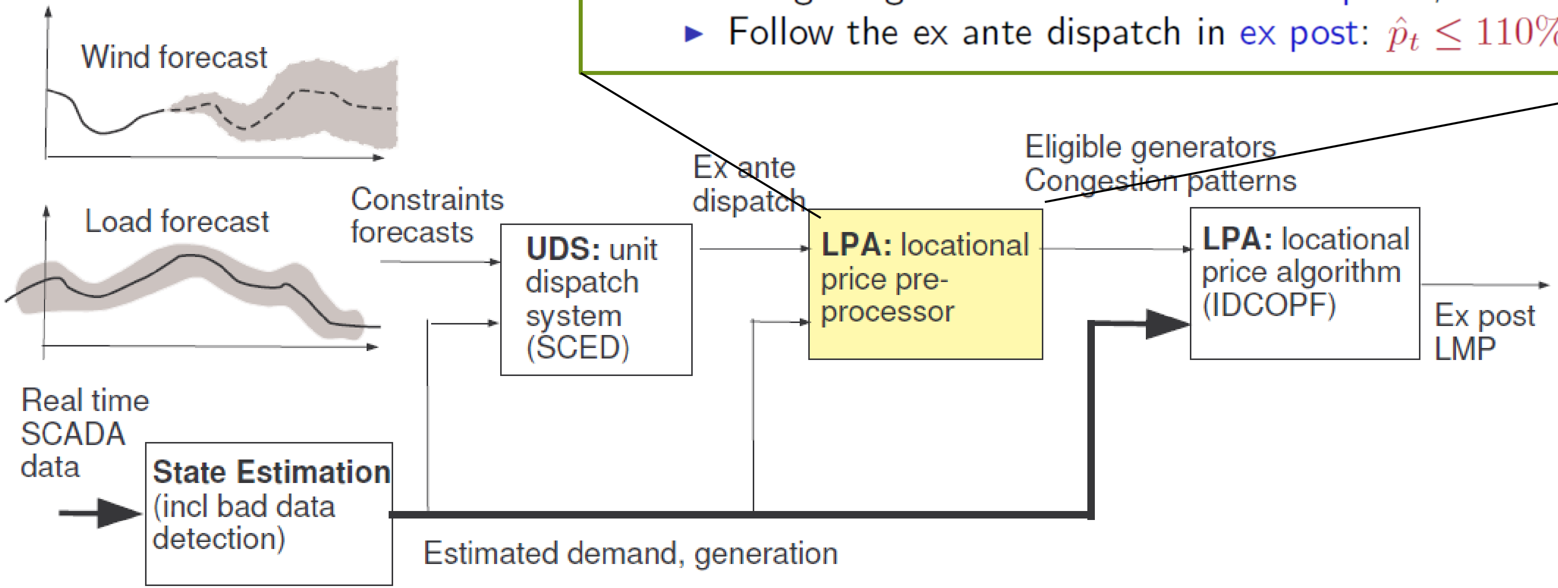


Ex ante dispatch

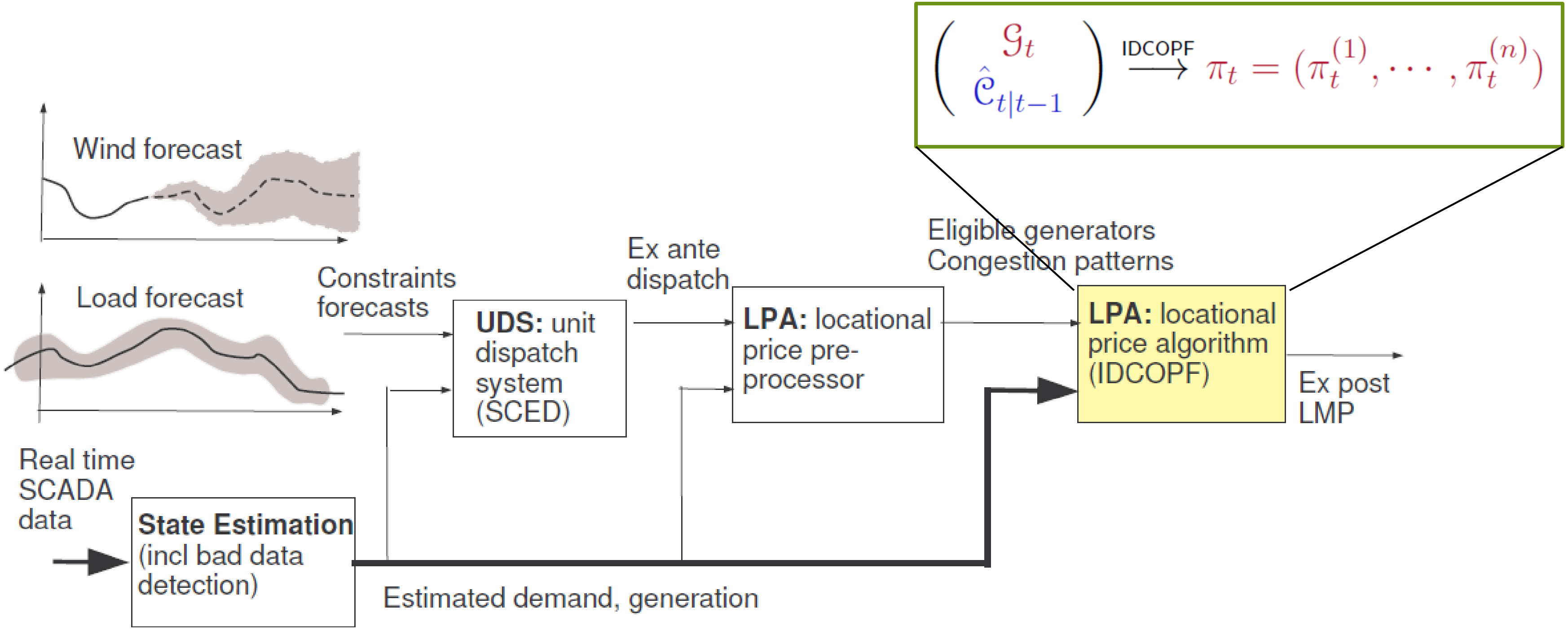


Ex post eligible generators

- Ex ante dispatch from SCOPF @ $t-1$: $\hat{p}_{t|t-1}$
- Ex post generation from state estimation @ t : $\hat{p}_t$
- Determine the set of eligible (flexible) generators $\mathcal{G}_t$:
  - Marginal generator in the ex ante dispatch;
  - Follow the ex ante dispatch in ex post: $\hat{p}_t \leq 110\% \times \hat{p}_{t|t-1}$



Ex post LMP



$\begin{pmatrix} \mathcal{G}_t \\ \hat{\mathcal{C}}_{t|t-1} \end{pmatrix} \xrightarrow{\text{IDCOPF}} \pi_t = (\pi_t^{(1)}, \cdots, \pi_t^{(n)})$
Wind forecast
Load forecast
Constraints forecasts
UDS: unit dispatch system (SCED)
Ex ante dispatch
LPA: locational price pre-processor
Eligible generators Congestion patterns
LPA: locational price algorithm (IDCOPF)
Ex post LMP
Real time SCADA data
State Estimation (incl bad data detection)
Estimated demand, generation

Ex post LMP via IDCOPF

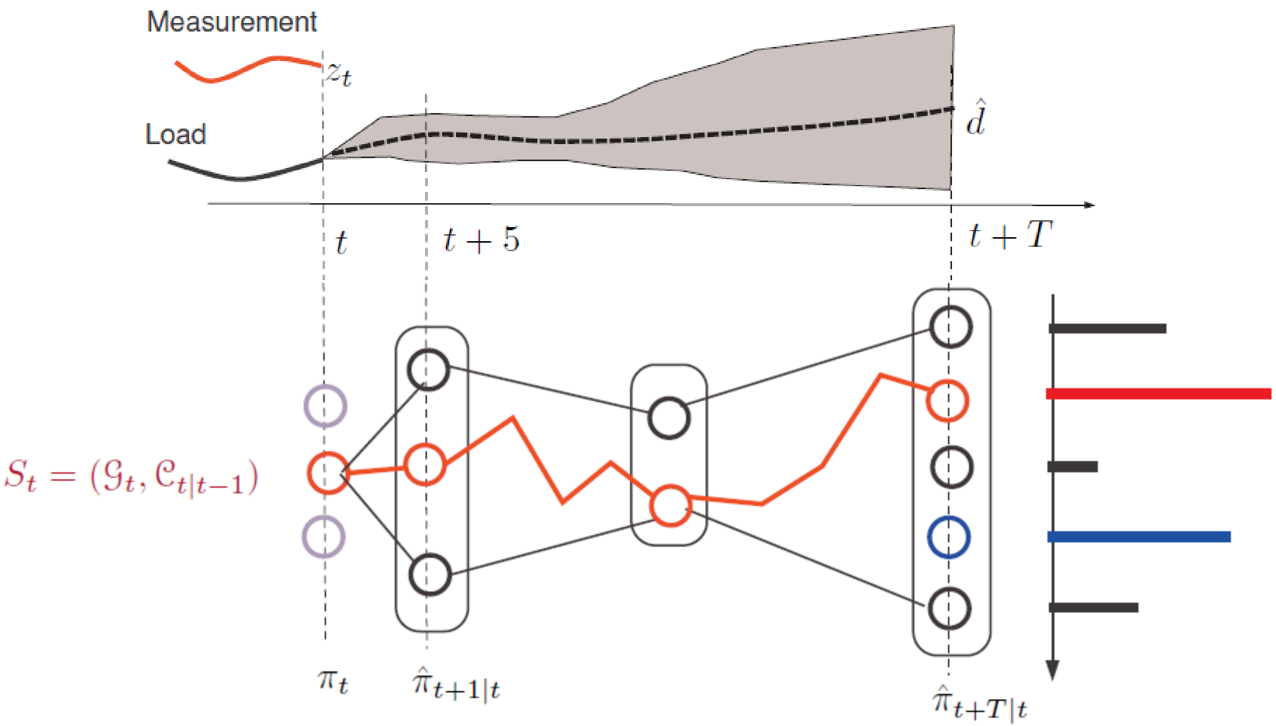
- Obtain the set of eligible generators $\mathcal{G}_t$ and the ex ante congestion $\hat{\mathcal{C}}_{t|t-1}$
- Compute an incremental DC OPF (Idcopf):

$$\begin{array}{llll} \text{minimize} & \sum_{i\in\mathcal{G}_t} c_i \Delta p_i & & \\ \text{subjcet to} & \sum_i \Delta p_i = 0 & & (\hat{\lambda}) \\ & \Delta p_{\text{min}} \leq \Delta p_i \leq \Delta p_{\text{max}} & i \in \mathcal{G}_t; & \\ & \sum_{i\in\mathcal{G}_t} A_{ki} \Delta p_i \leq 0; & k \in \mathcal{C}_{t|t-1} & (\hat{\mu}_k) \end{array}$$

- The vector LMP is given by

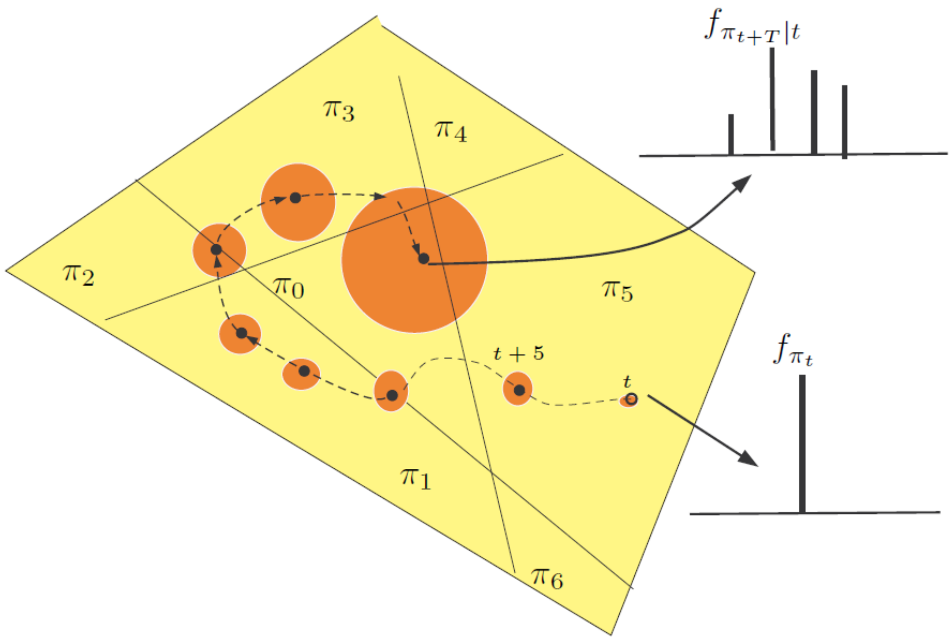
$$\hat{\pi} = \hat{\lambda}^* \mathbf{1} - A\hat{\mu} \triangleq \text{Idcopf}(\mathcal{G}_t, \mathcal{C}_{t|t-1})$$

Information structure and Markov chain



Probabilistic forecasting

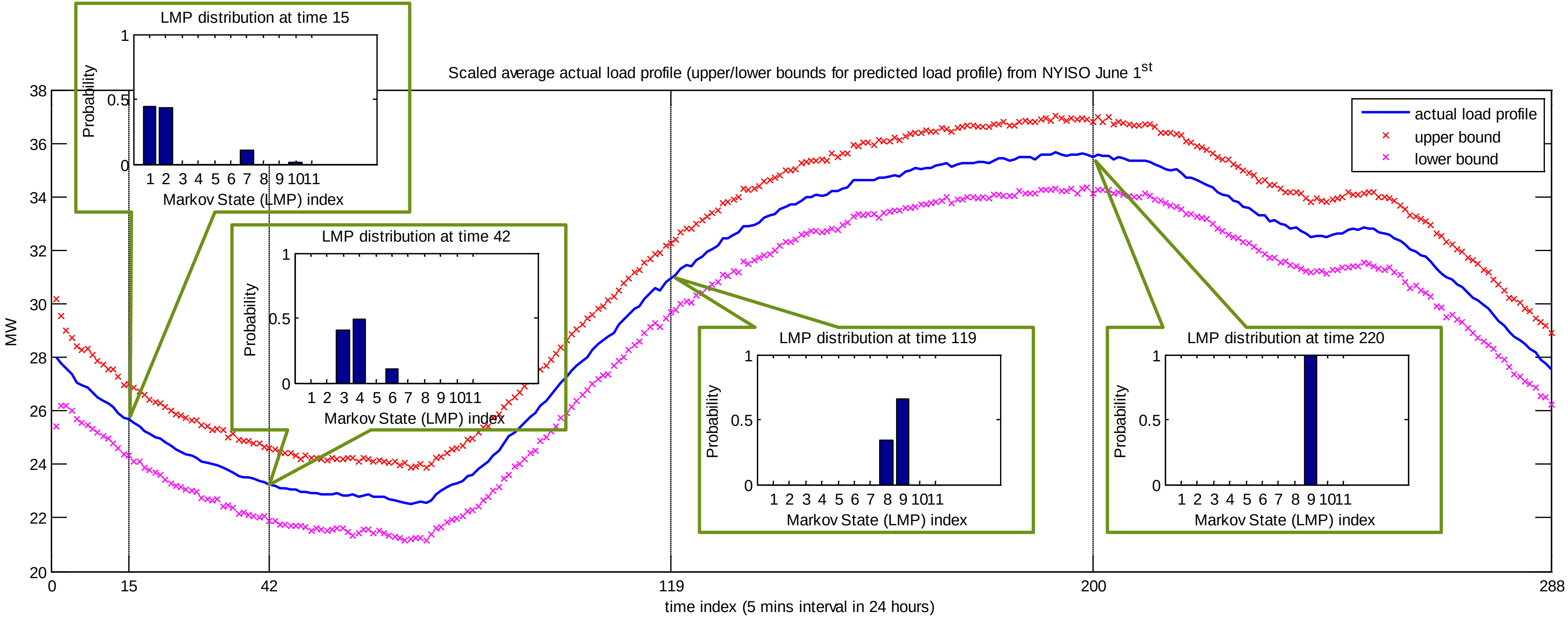
$$f_{\pi_{t+T}|t} = \delta_{\pi_t} \times P_t \times P_{t+5} \times \cdots \times P_{t+T-1}$$



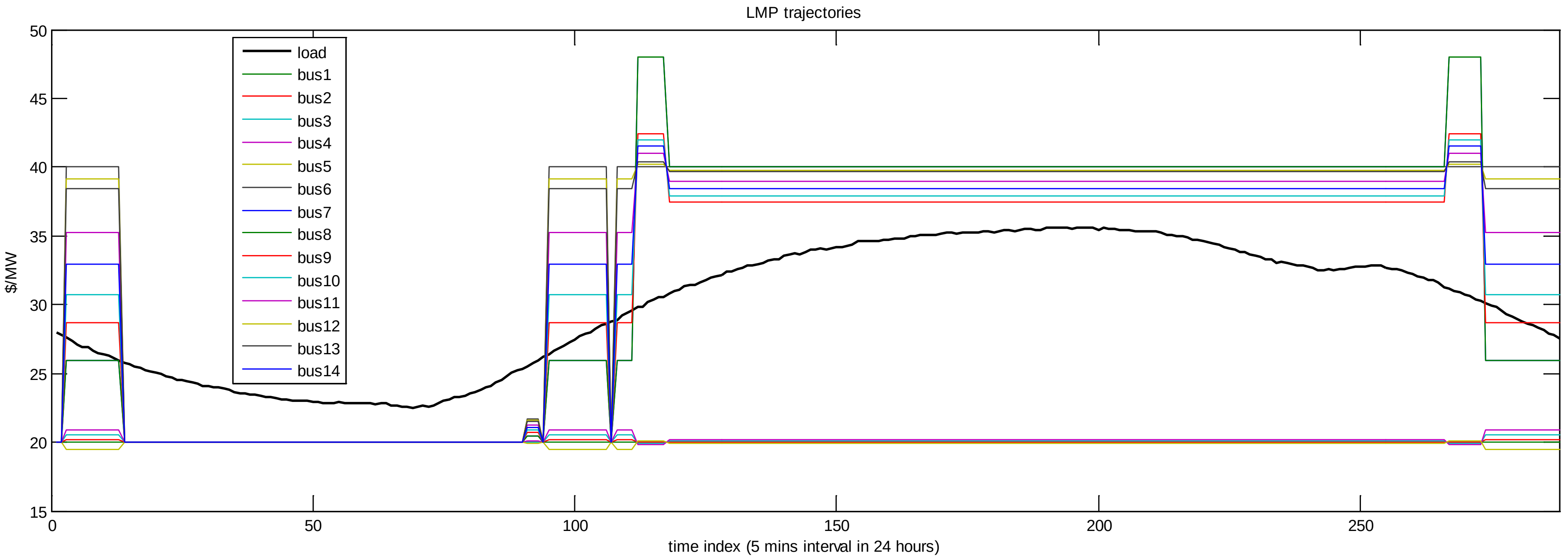
Outline

- Motivation
- Probabilistic forecasting of real-time LMP
- Preliminary simulations
  - IEEE 16 bus with NYISO sample load profiles
  - Varying load errors and correlations
  - Monte Carlo techniques to estimate transition matrices
- Summary and future work

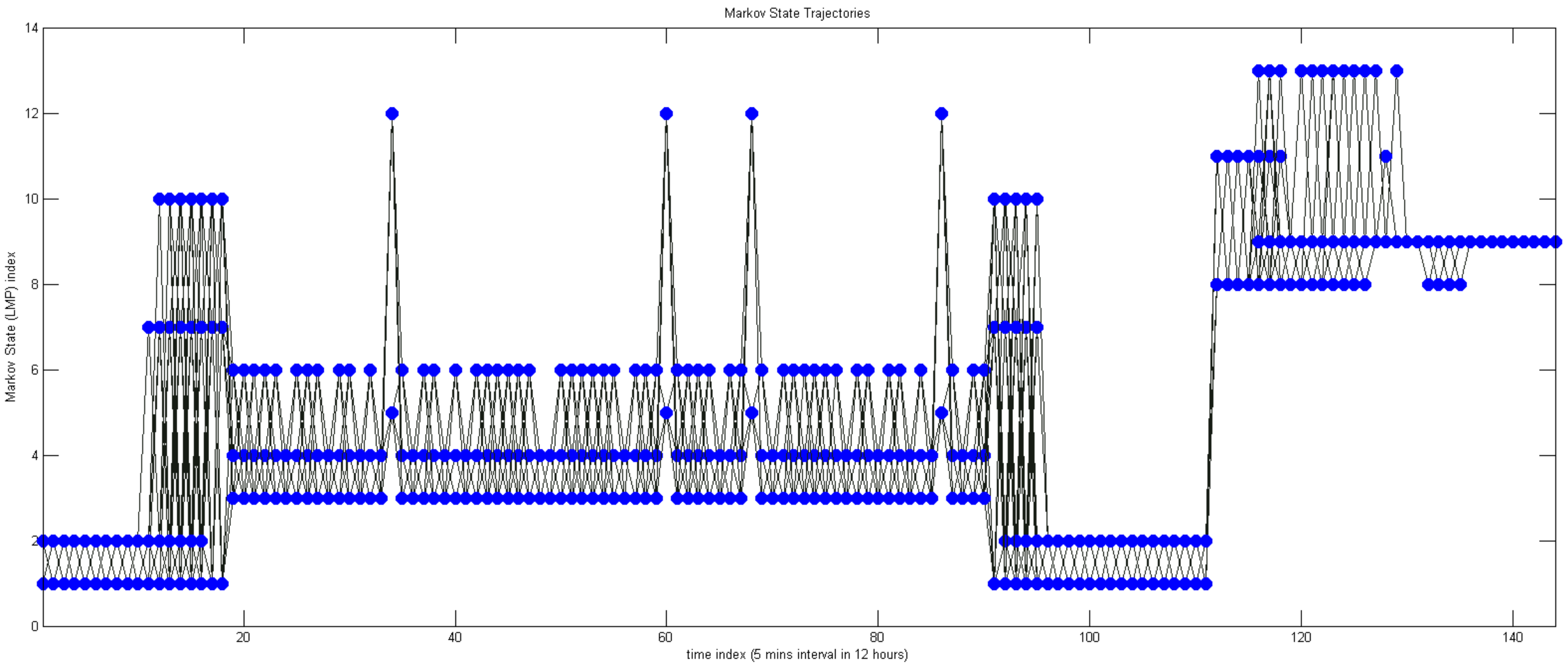
Load and price distributions



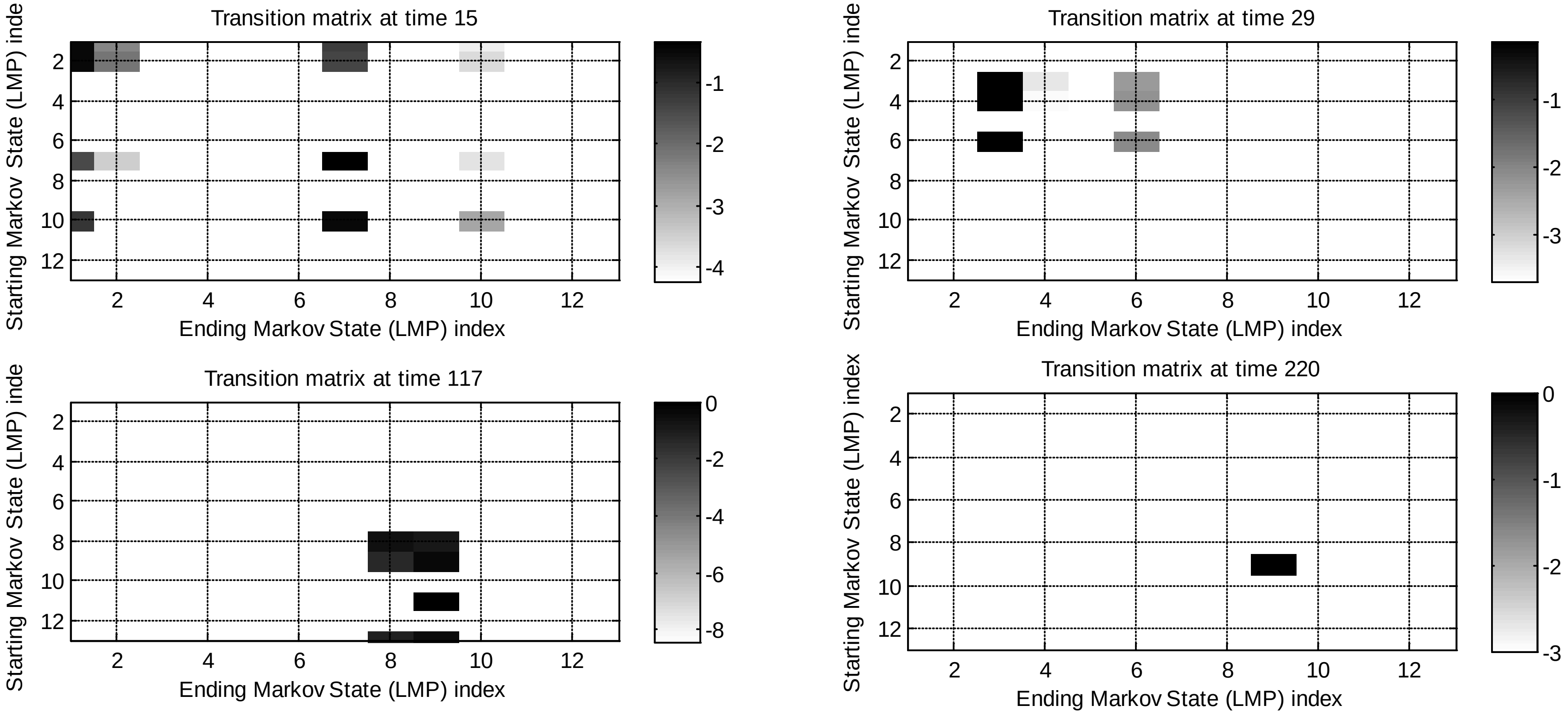
LMP trajectories



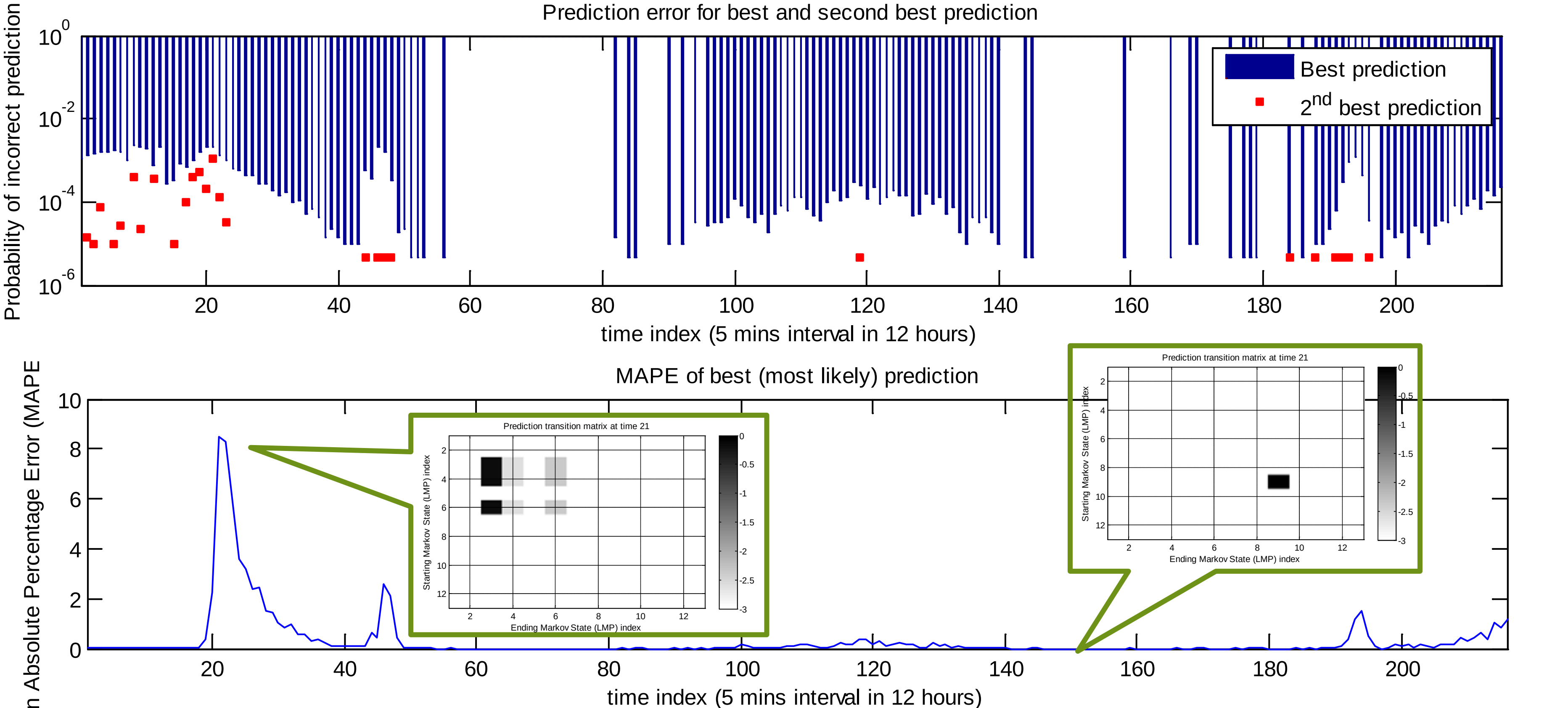
Time inhomogeneous Markov Chain



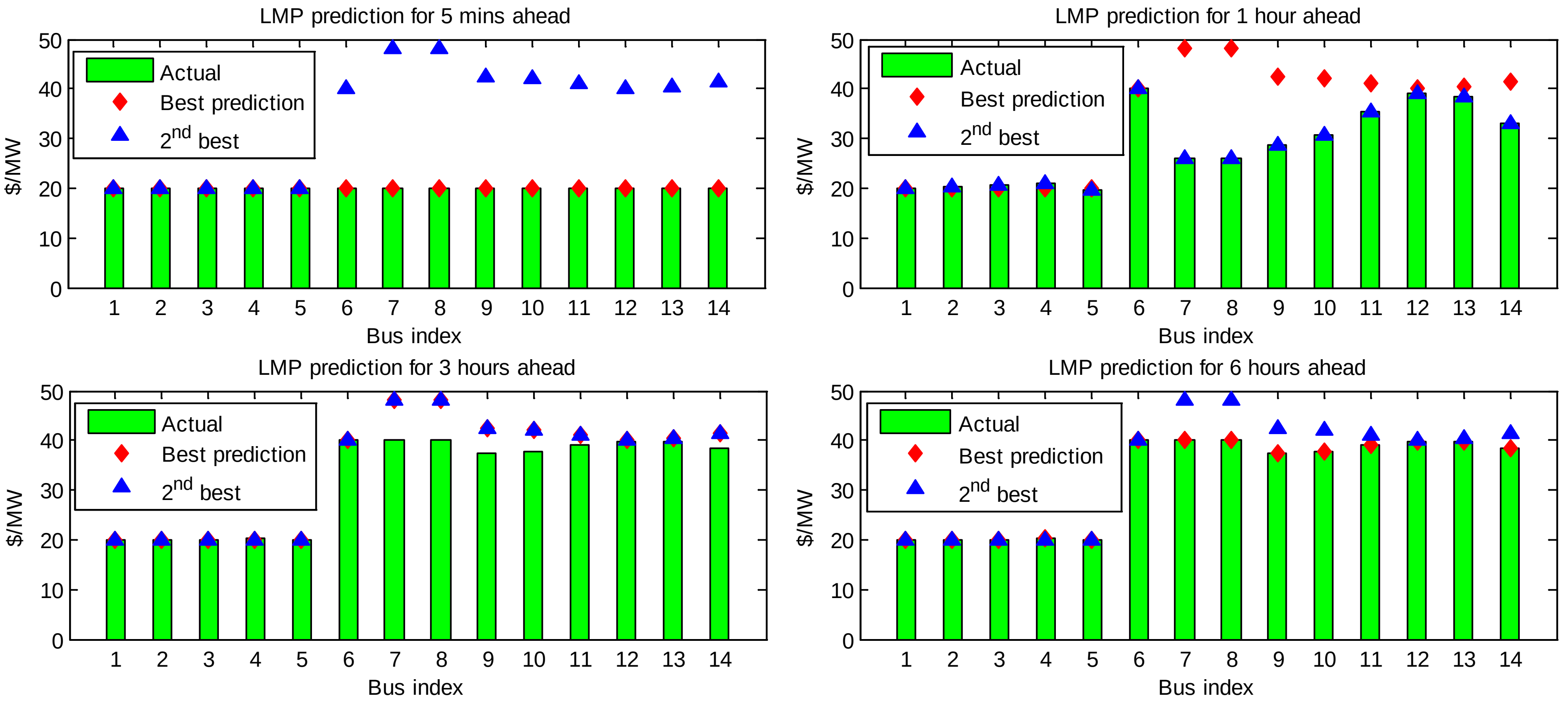
Transition Matrices



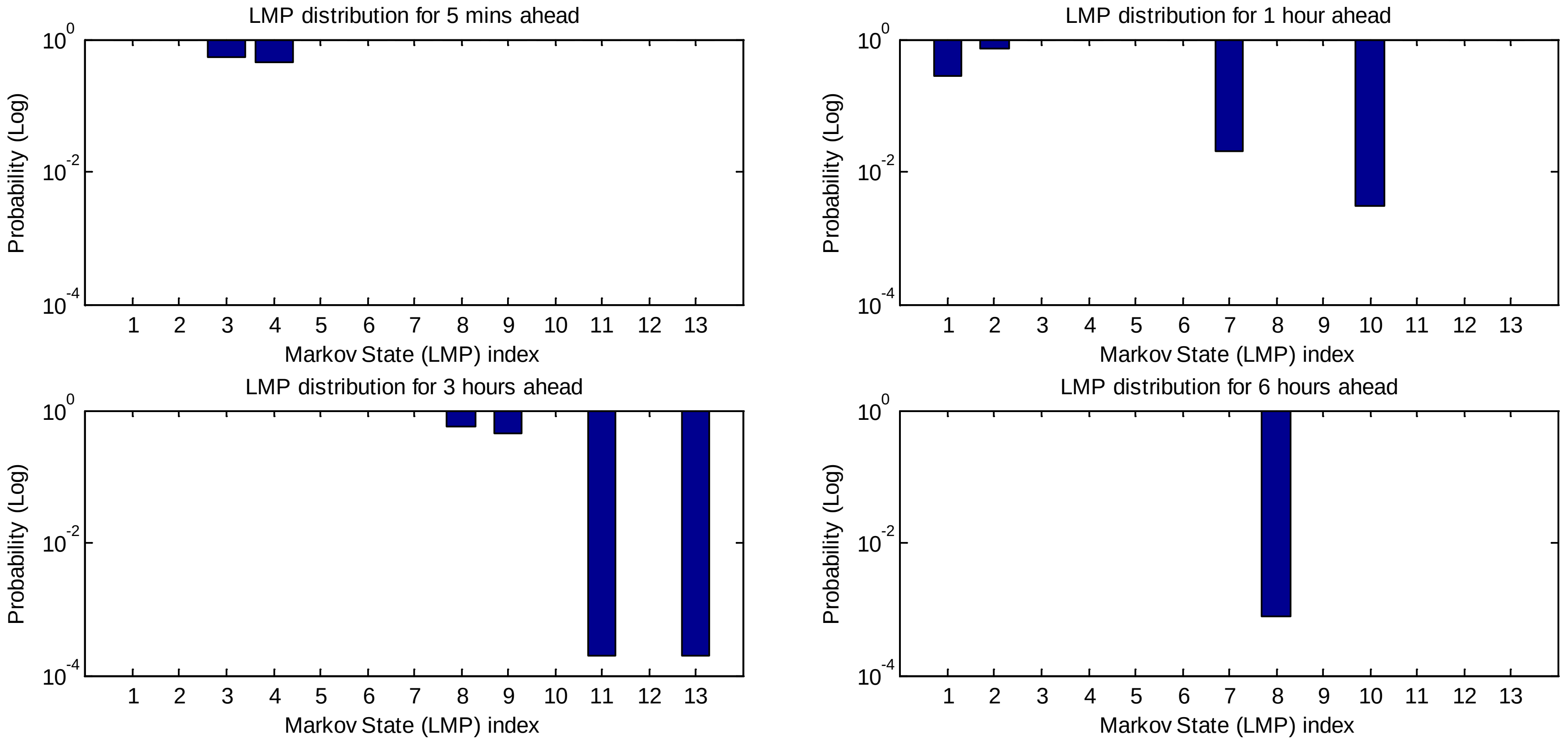
6 hours ahead prediction error



LMP prediction for different time horizon



LMP distribution for different prediction horizons



Planned research

- Forecasting techniques
  - Incorporating load models
  - More sophisticated Monte Carlo techniques (important sampling, MCMC)
  - Generator behavior models
  - Forecasting with renewable sources
  - Demand response
- Extensive simulation studies and comparisons

DATA QUALITY AND ITS EFFECTS ON MARKET OPERATIONS

DENY OF SERVICE ATTACK ON REAL-TIME ELECTRICITY MARKET

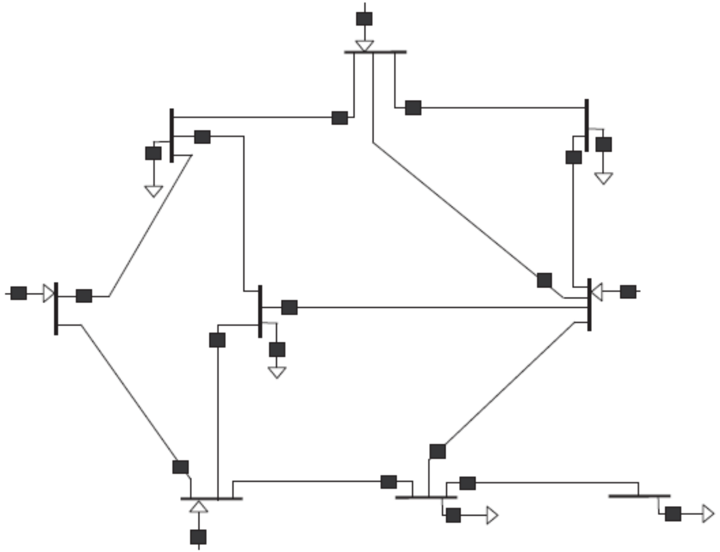
Lang Tong (PI), Robert J. Thomas, and Jinsub Kim

Cornell University

Presented at CERTS Review

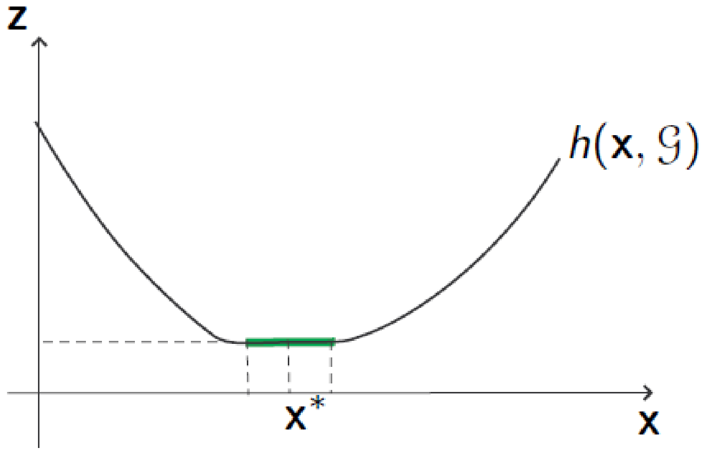
August 7-8, 2013

Observability

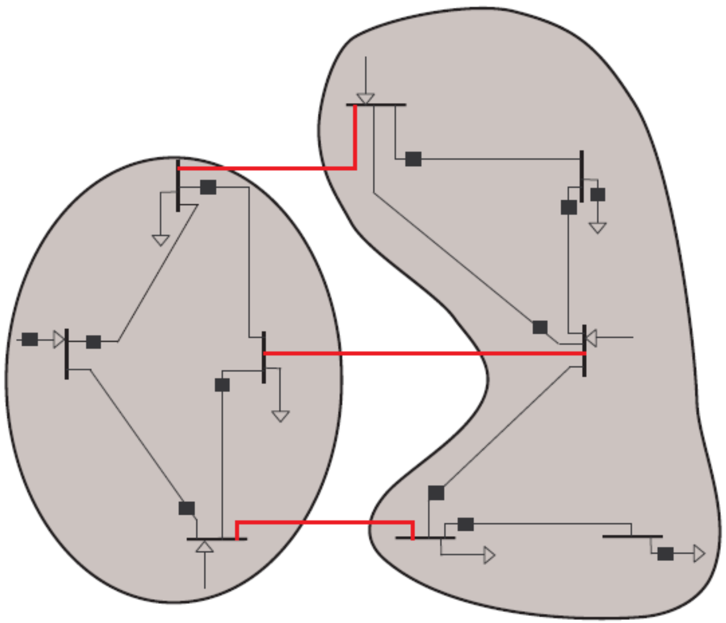


- Power system as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
  - System state x: voltage phasors at buses (vertices)
  - Observation equation: $z = h(x, \mathcal{G})$

- Locally unobservable at $\mathbf{x}^*$:



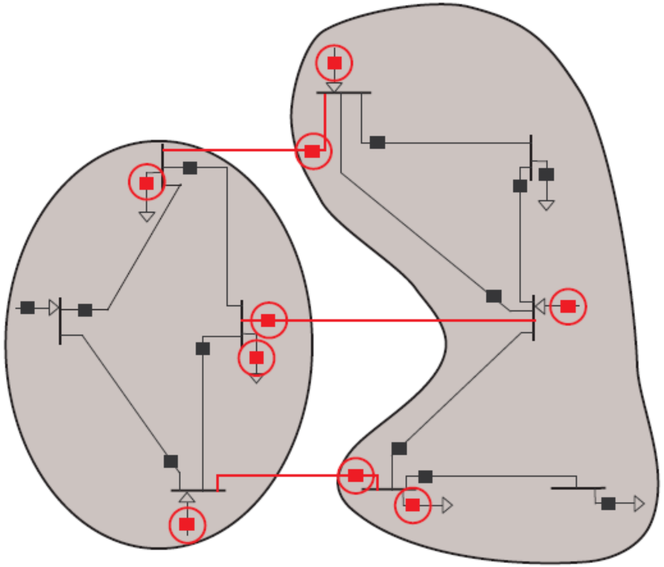
A graph theoretic condition (KCD'80)



Theorem

A network is observable if and only if $\exists$ *a spanning tree with an assigned meter on each edge.*

Undetectable attack (KJTT'11)

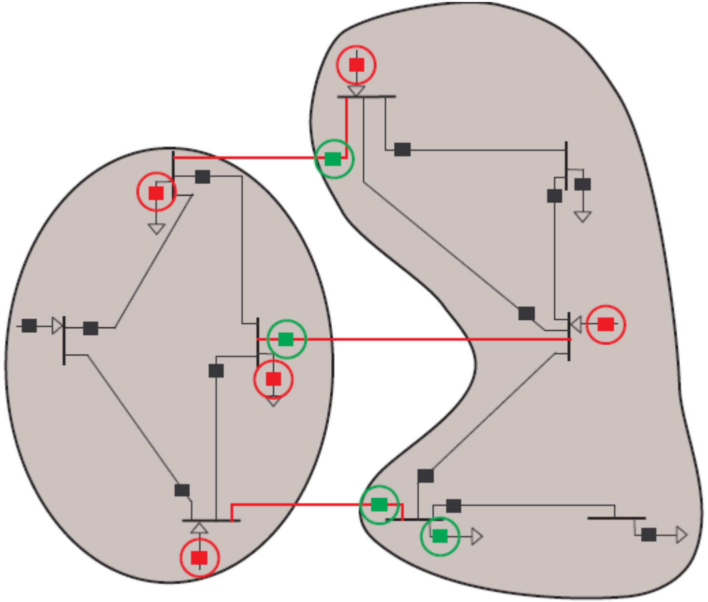


Theorem

The attacker can make the power system unobservable by controlling a set of meters associated with a cut without being detected by the control center.

- The minimum number of meters accessible to the attacker is $\Theta(|\text{min-cut}|)$.

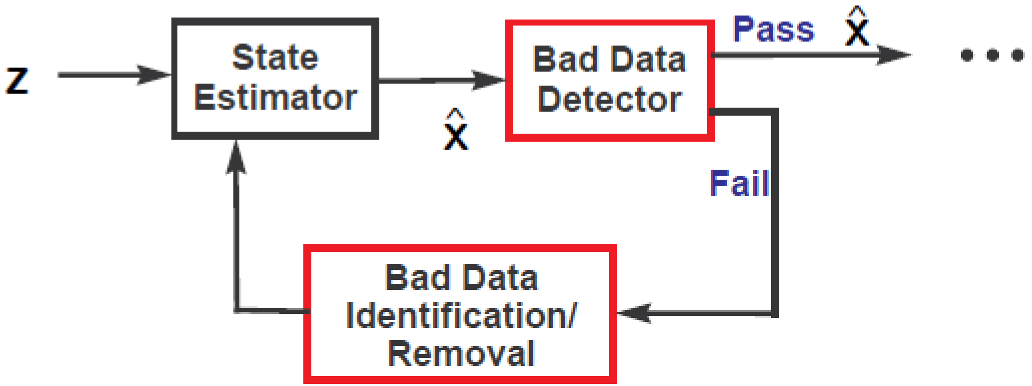
Only half of the meters are needed!



Framing attack:

- Let $\mathcal{J}$ be a set of meters on a cut with partition $\mathcal{J} = \mathcal{J}_1 \bigcup \mathcal{J}_2$.
- The attacker injects bad data on $\mathcal{J}_1$ such that the control center (i) detects bad data; (ii) identifies bad data occur at $\mathcal{J}_2$.

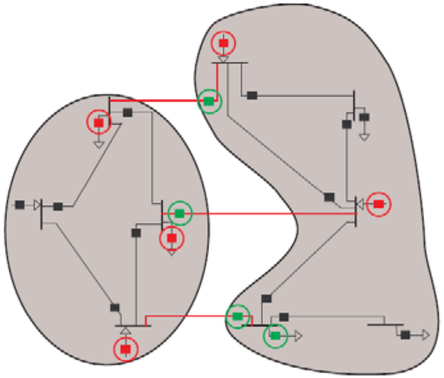
Bad data identification and removal

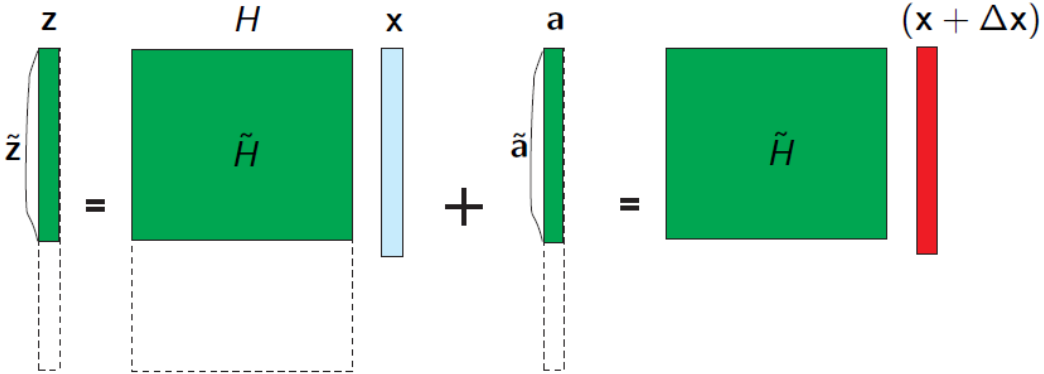


- State estimate: $\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} (\mathbf{z} - h(\mathbf{x}))^T \Sigma^{-1} (\mathbf{z} - h(\mathbf{x}))$
- Residue: $\mathbf{r} \triangleq \mathbf{z} - h(\hat{\mathbf{x}})$; $\quad \mathbf{r}^N \triangleq \Omega^{-1}\mathbf{r}$.
- Bad data identification and removal:
  - If $\mathbf{r}^T \Sigma^{-1} \mathbf{r} \leq \tau$, accept $\hat{\mathbf{x}}$;
  - If $\mathbf{r}^T \Sigma^{-1} \mathbf{r} > \tau$, remove the meter i with the largest $|r_i^N|$.

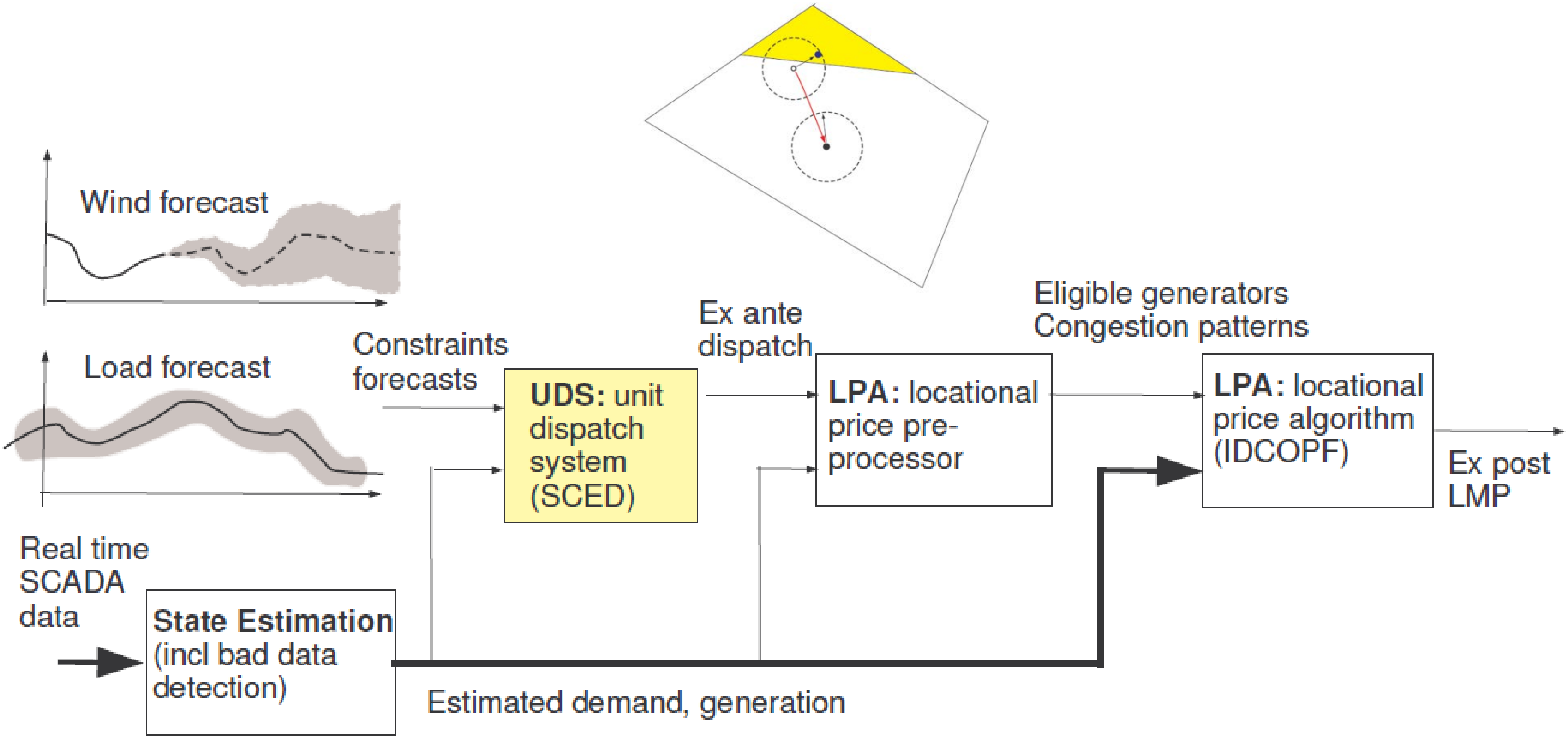
Framing attack via QCQP

maximize	$\mathbb{E}\{\sum_{i\in\mathcal{I}_T}(r_i^N)^2\} = \|R\mathbf{a}\|_2^2$	maximizing residue energy of target meters.
subject to	$\|\mathbf{a}\|_2 = 1, \mathbf{a} \in \mathcal{A}$	find target direction
	$\tilde{\mathbf{a}} \in \text{Col}(\tilde{H})$	$(\tilde{\mathbf{a}} = \tilde{H}\Delta\mathbf{x})$.

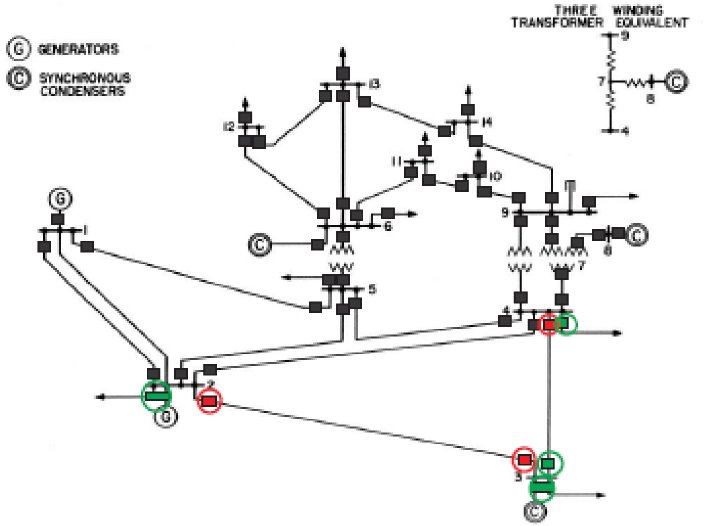


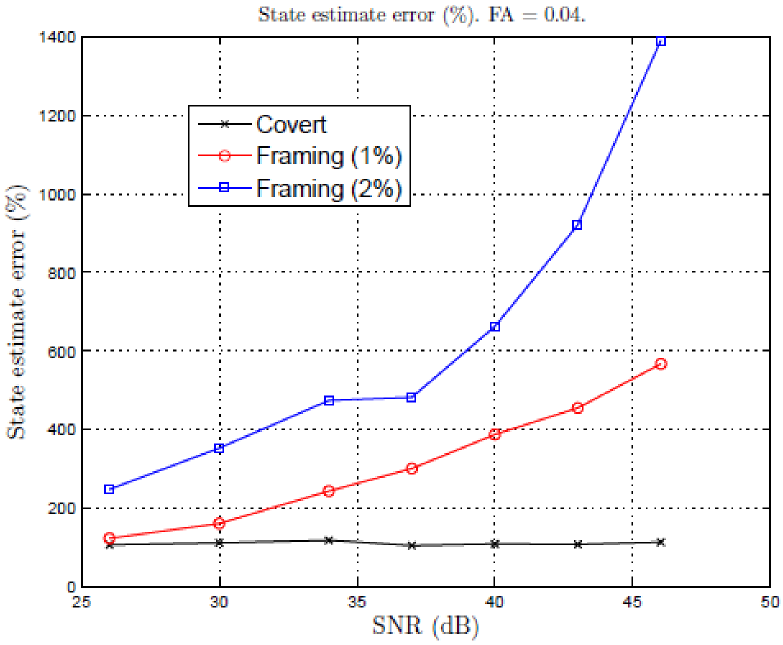


DoS attack on real-time operations



Simulation examples (AC)





Conclusion

- Attacks on cyber physical systems (power systems) are real:
  - State attacks
  - Topology attacks
  - Dispatch attacks
- A key step toward protection is deploy authentication mechanisms.