NILM Beyond Event Detection

J. Zico Kolter
MIT CSAIL

NILM Workshop - 5/7/12

In collaboration with: Tommi Jaakkola

Origins

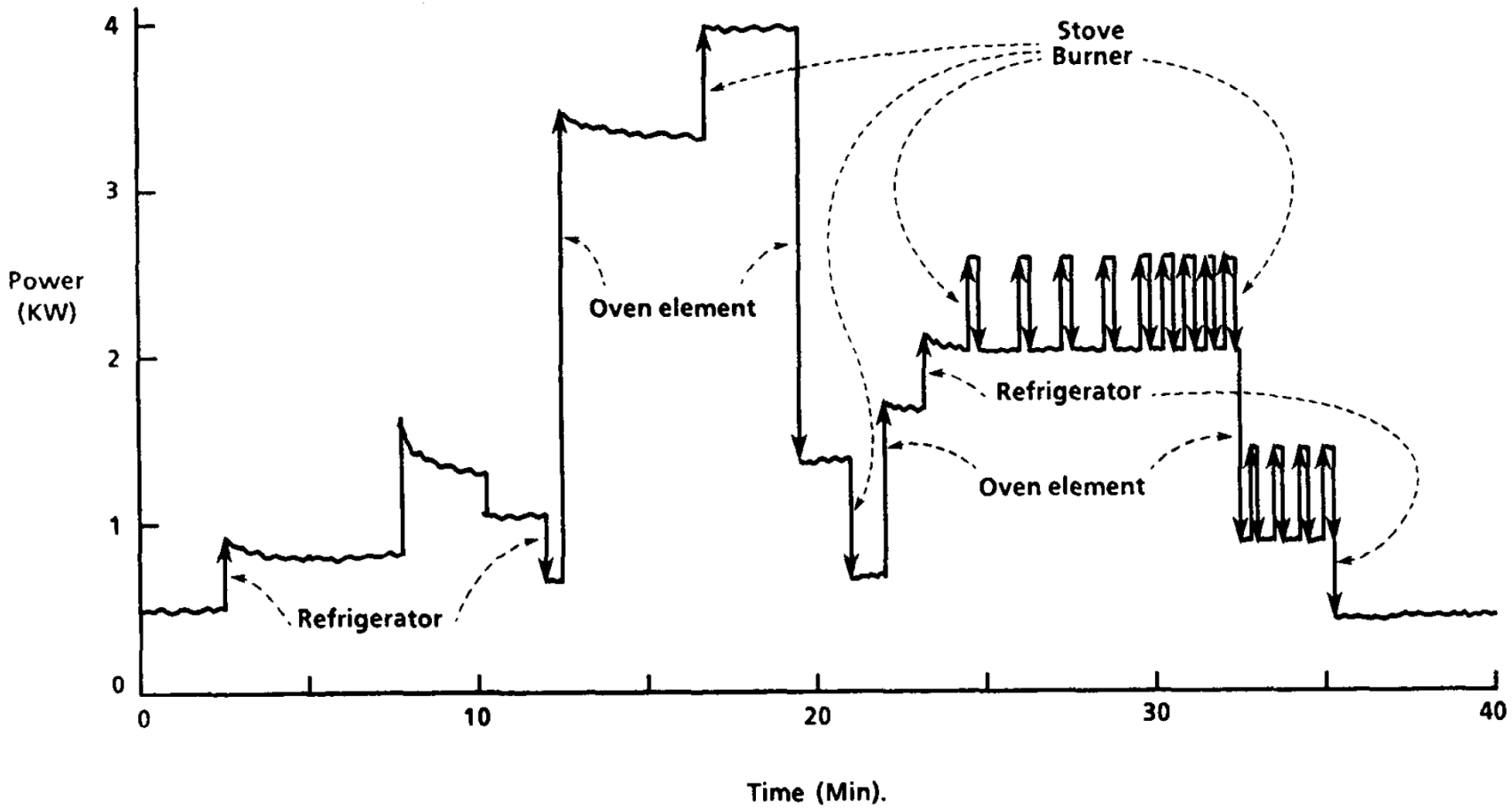


Figure from G.W. Hart, "Nonintrusive Application Load Monitoring." Proceedings of the IEEE, Vol 80 (12), 1992.

Origins

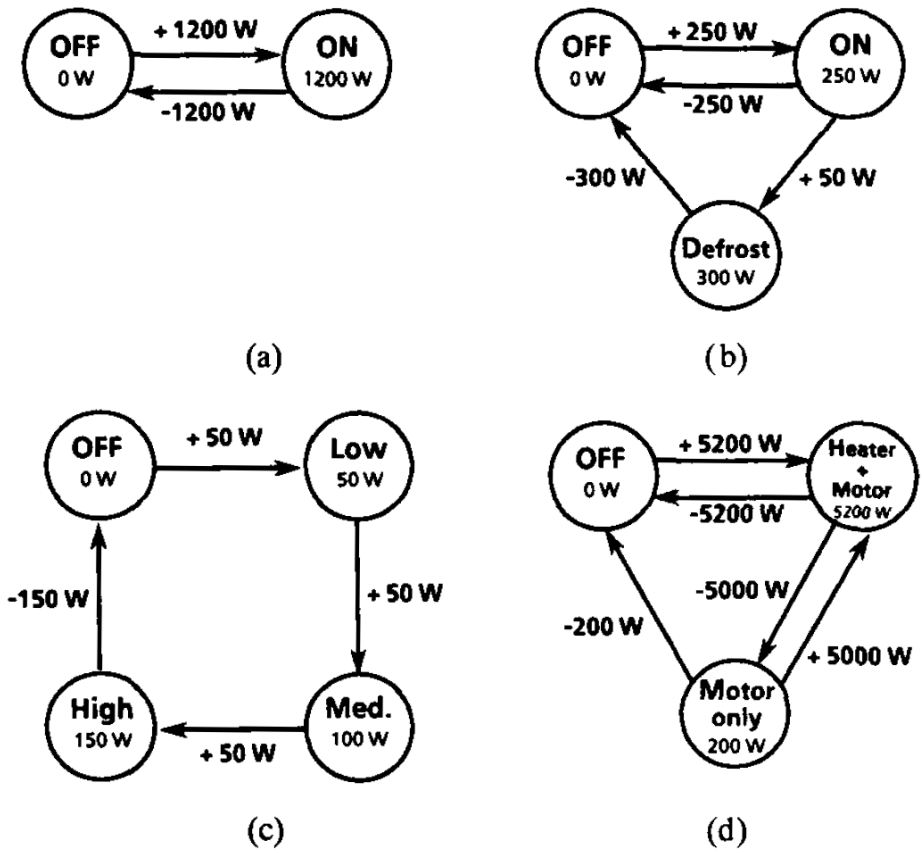


Fig. 4. Finite-state appliance models: (a) generic 1200 W two-state appliance, e.g., toaster; (b) refrigerator with defrost state; (c) "three-Way" lamp; (d) clothes dryer.

Figure from G.W. Hart, "Nonintrusive Application Load Monitoring." Proceedings of the IEEE, Vol 80 (12), 1992.

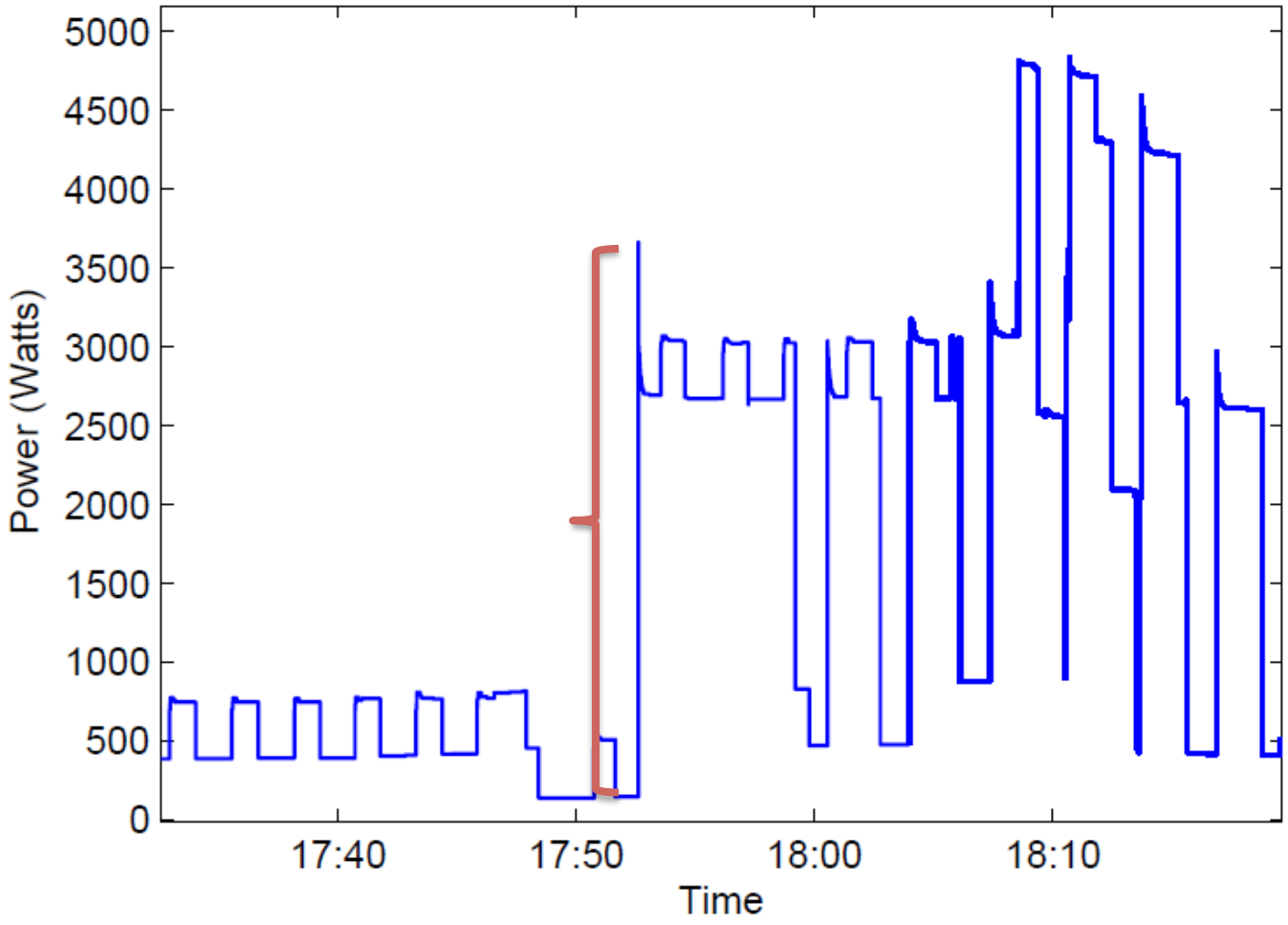
A. Bouloutas, G.W. Hart, M. Schwartz. "Two extensions of the Viterbi Algorithm." IEEE Transactions on Information Theory, Vol 37 (3), 1991

G.W. Hart, A. Bouloutas. "Correcting Dependent Errors in Sequences Generated by Finite-State Processes." IEEE Transactions on Information Theory, Vol 39 (4), 1991

Typical Approach

1. ***Specify*** model

- For each device: $\triangle P^{(i)}$



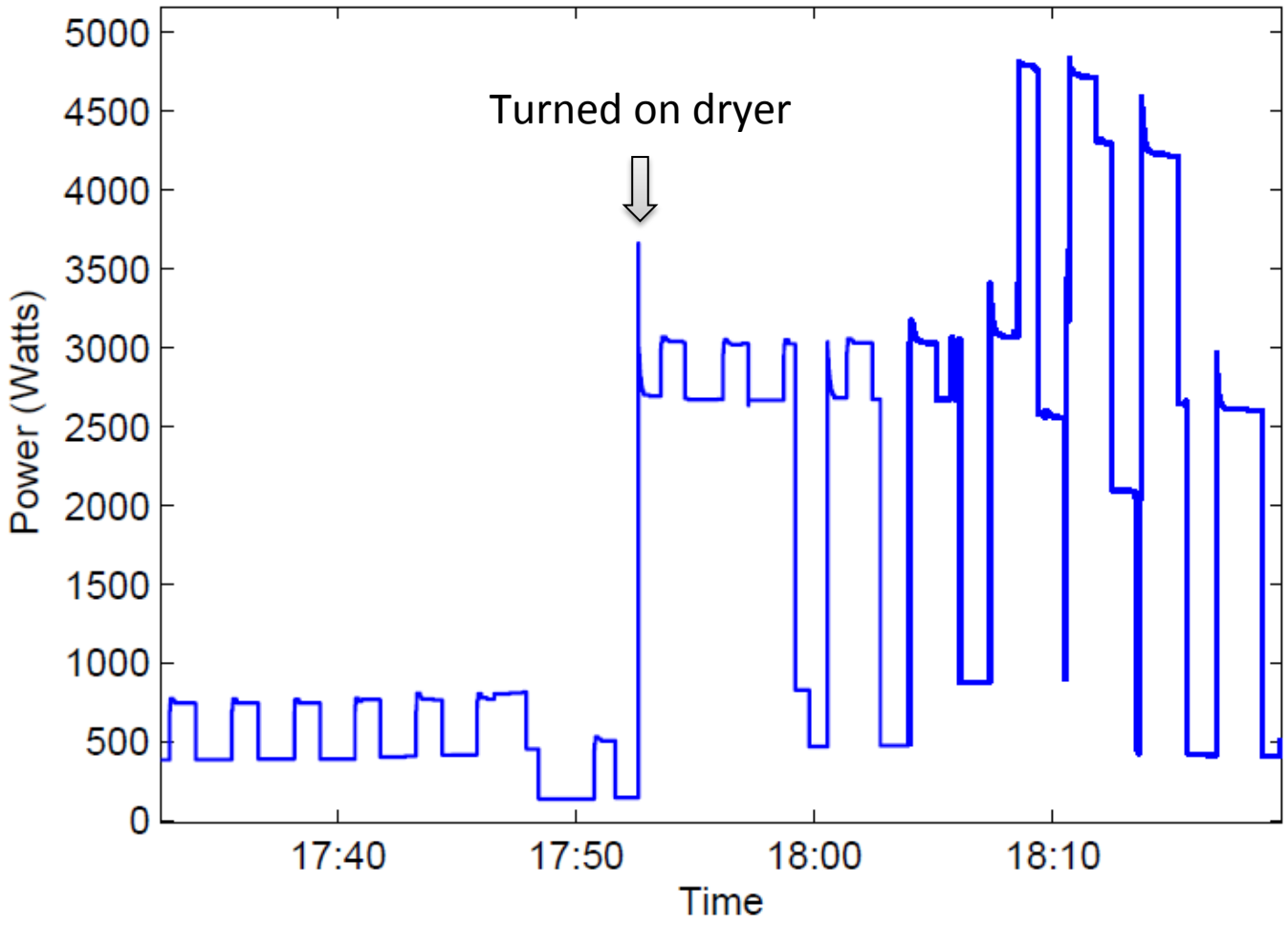
Typical Approach

1. ***Specify*** model

- For each device: $\triangle P^{(i)}$

2. ***Learning*** (supervised)

- Turn each device on/off, record power delta



Typical Approach

1. ***Specify*** model

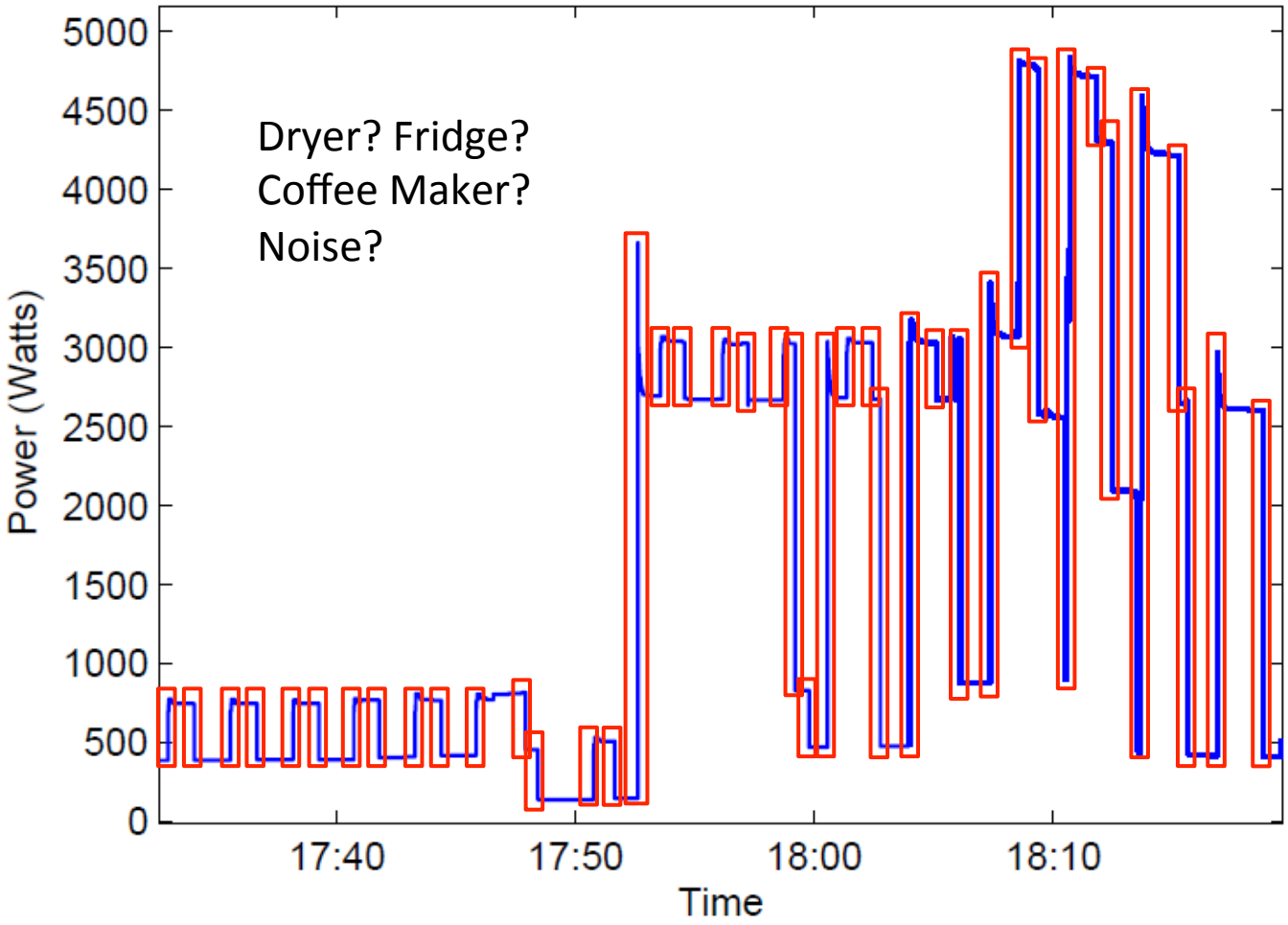
- For each device: $\triangle P^{(i)}$

2. ***Learning*** (supervised)

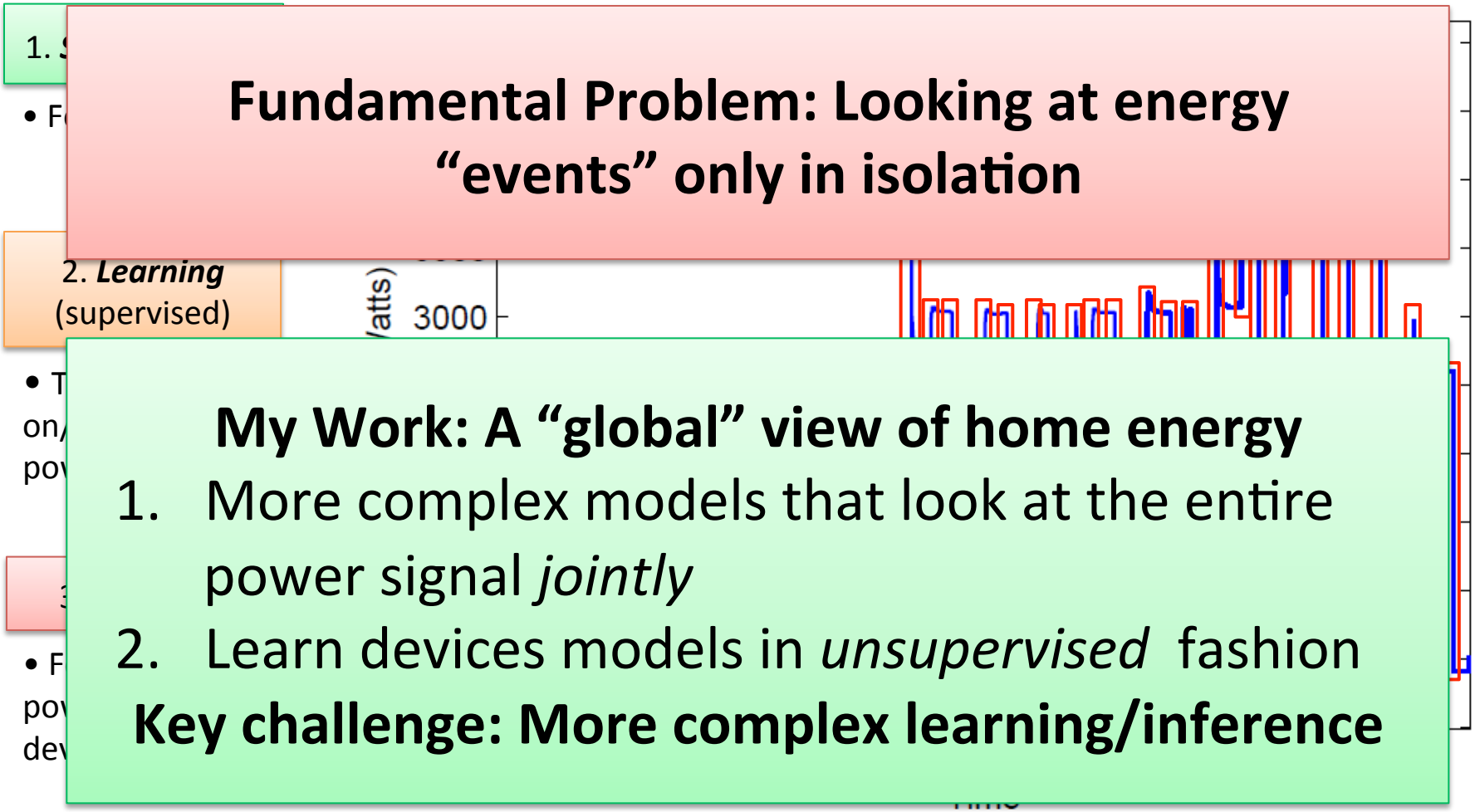
- Turn each device on/off, record power delta

3. ***Inference***

- For each change in power, find nearest device match



Typical Approach



Fundamental Problem: Looking at energy "events" only in isolation

My Work: A "global" view of home energy

1. More complex models that look at the entire power signal *jointly*
2. Learn devices models in *unsupervised* fashion

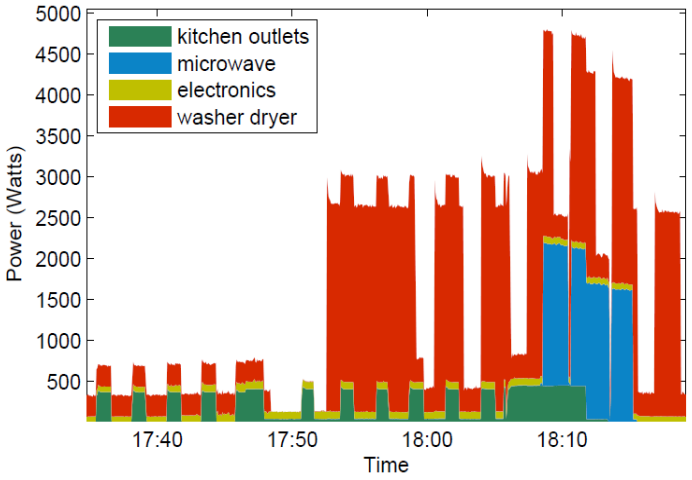
Key challenge: More complex learning/inference

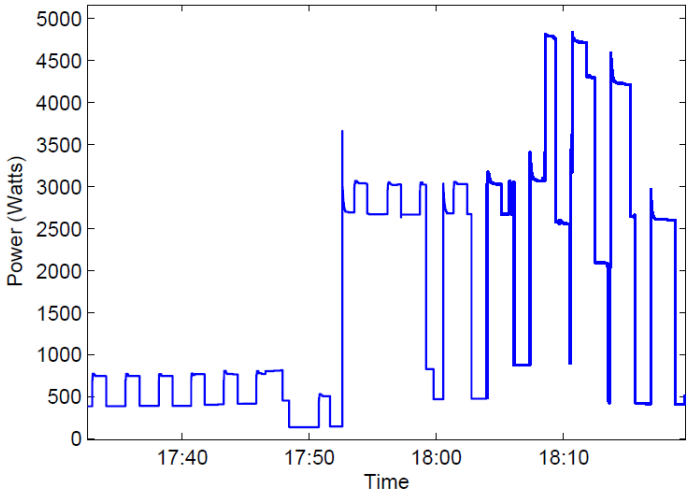
Problem Formulation

Find:
power consumptions
for each device

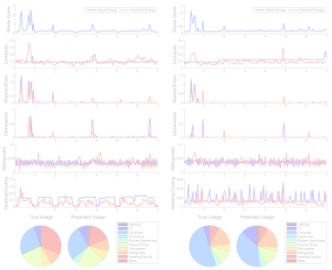
$$\begin{array}{cccccc} y_1^{(1)} & y_2^{(1)} & y_3^{(1)} & \cdots & y_T^{(1)} \\ y_1^{(2)} & y_2^{(2)} & y_3^{(2)} & \cdots & y_T^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ \hline \bar{y}_1 & \bar{y}_2 & \bar{y}_3 & \cdots & \bar{y}_T \end{array}$$

Given:
whole-home power
consumption over time





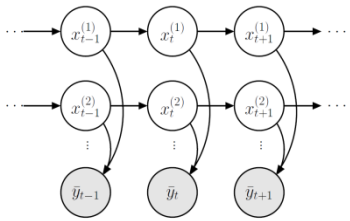
Work on Disaggregation



- Initial work using discriminative sparse coding, hourly power data
  [Kolter, Batra and Ng, NIPS 2010]



- Data collection, compression, and a public dataset
  [Kolter and Johnson, KDDSUST 2011]



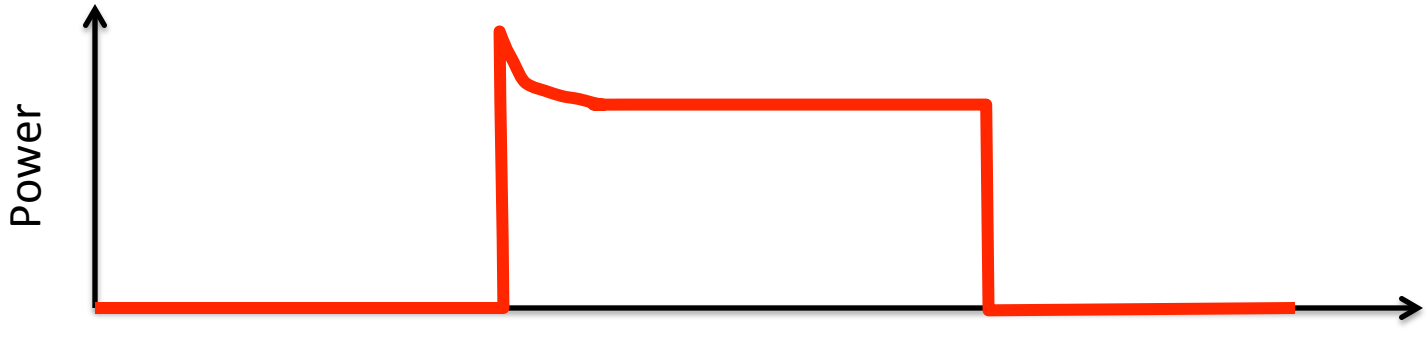
- Factorial HMM modeling and efficient approximate inference
  [Kolter and Jaakkola, AISTATS 2012]

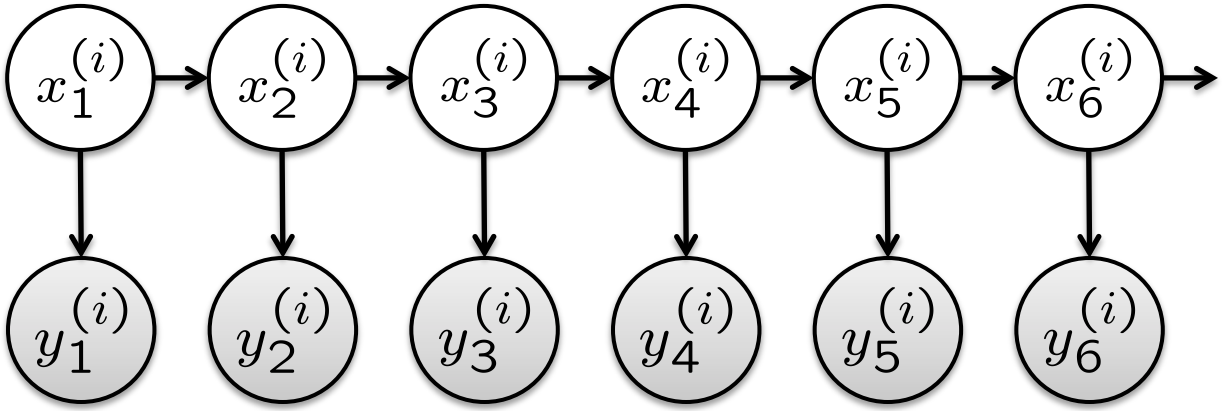
Devices as Hidden Markov Models



1. ***Specify*** model

Devices as Hidden Markov Models



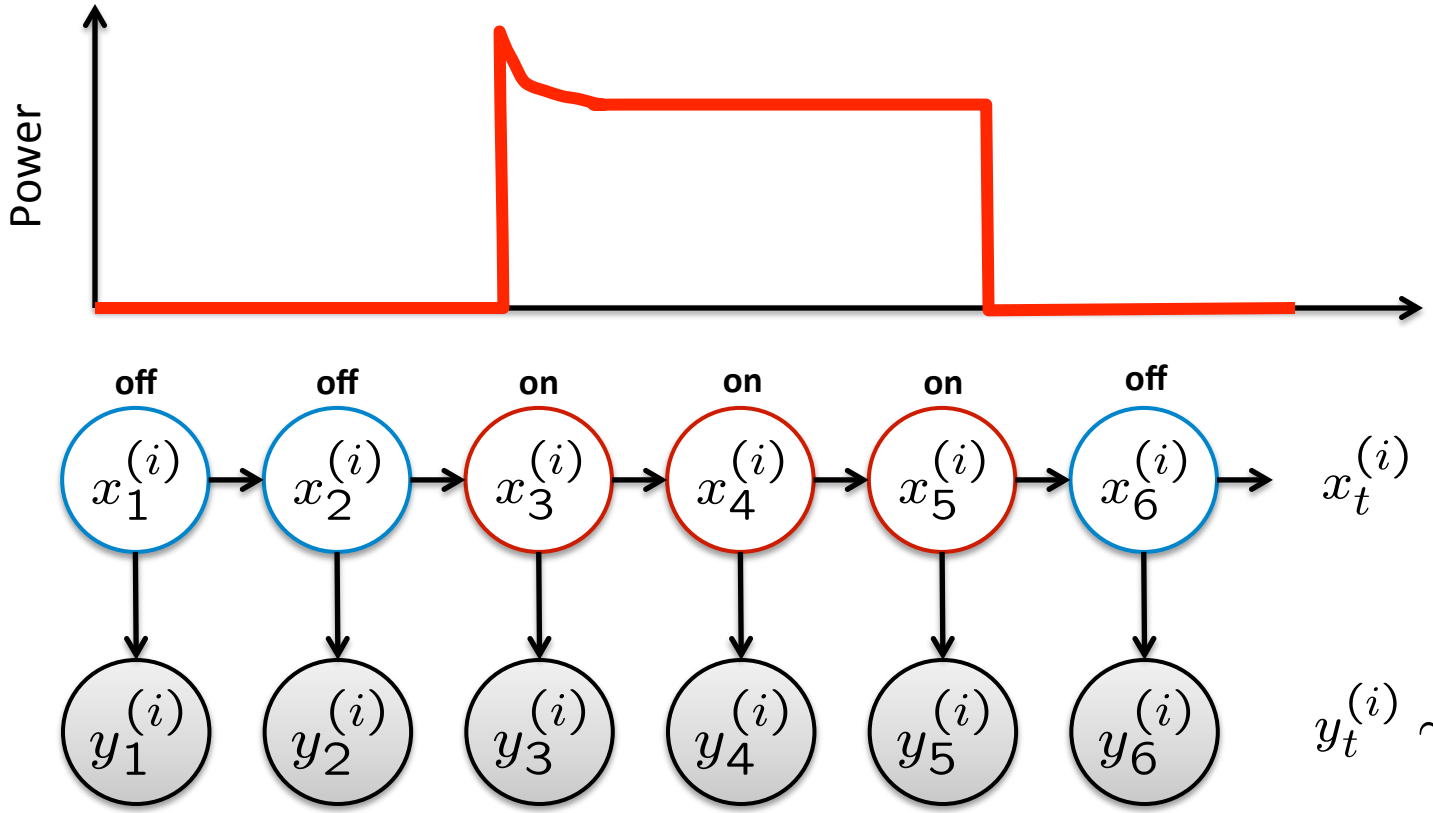


$$x_t^{(i)} \in \{1, \ldots, m_i\}$$

$$y_t^{(i)} \sim \mathcal{N}(\mu_{x_t^{(i)}}^{(i)}, \Sigma)$$

1. ***Specify*** model

Devices as Hidden Markov Models

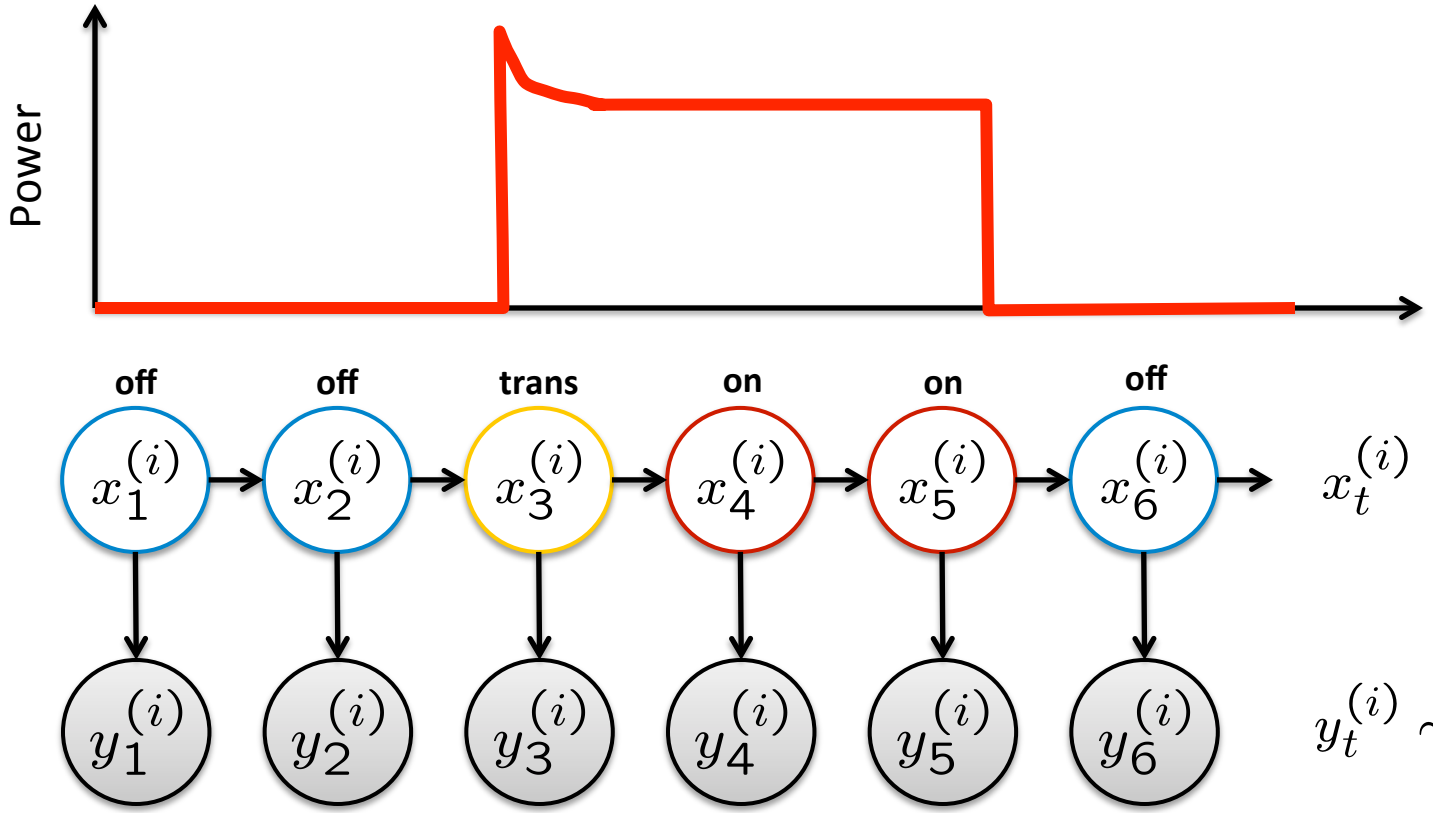


$$x_t^{(i)} \in \{1, \ldots, m_i\}$$

$$y_t^{(i)} \sim \mathcal{N}(\mu_{x_t^{(i)}}^{(i)}, \Sigma)$$

1. ***Specify*** model

Devices as Hidden Markov Models

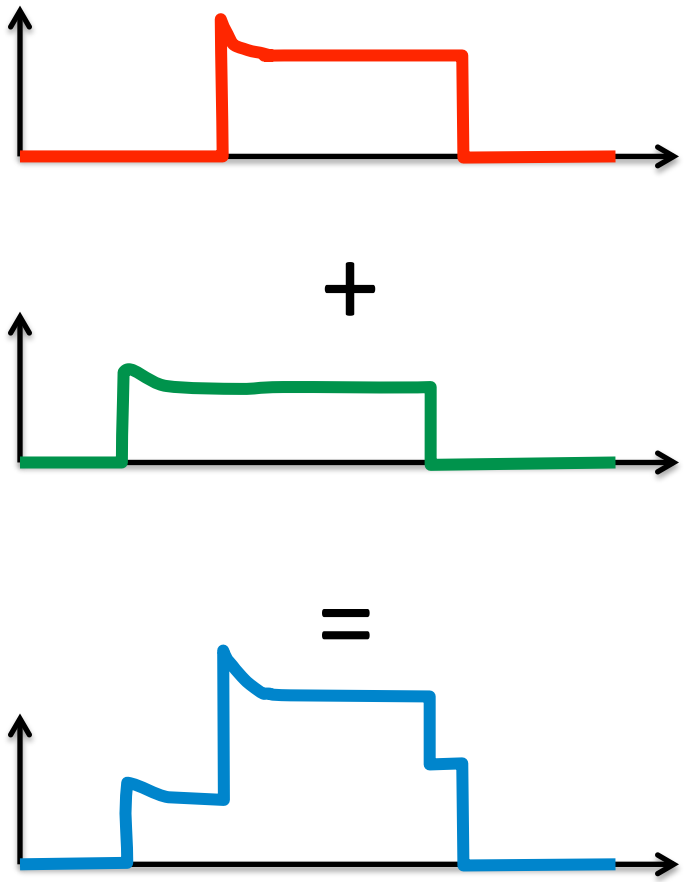


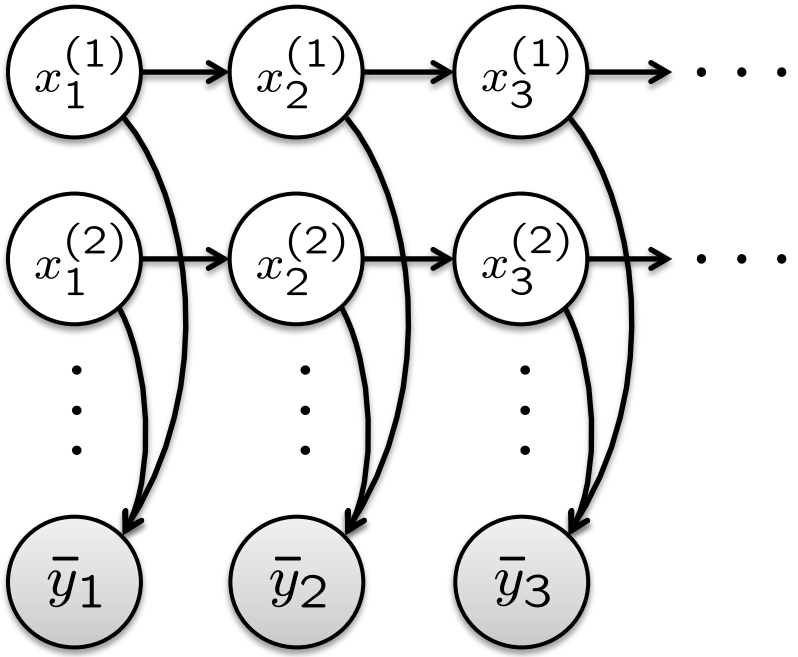
$$x_t^{(i)} \in \{1, \ldots, m_i\}$$

$$y_t^{(i)} \sim \mathcal{N}(\mu_{x_t^{(i)}}^{(i)}, \Sigma)$$

1. ***Specify*** model

Factorial HMMs





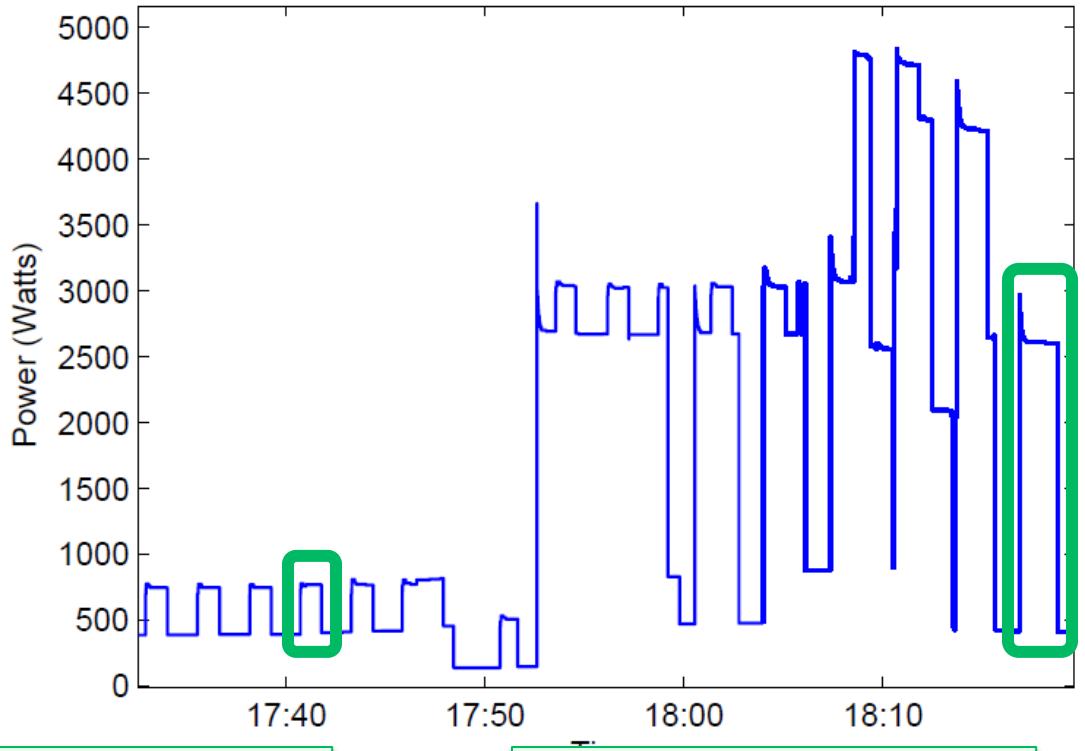
$$\bar{y}_t | x_t^{(1:N)} \sim \mathcal{N}\left(\sum_{i=1}^{N} \mu_{x_t^{(i)}}^{(i)}, \Sigma\right)$$

[Ghahramani and Jordan, 1997]

1. ***Specify*** model

Learning Device Models

Intuition: we do sometimes observe devices events in isolation

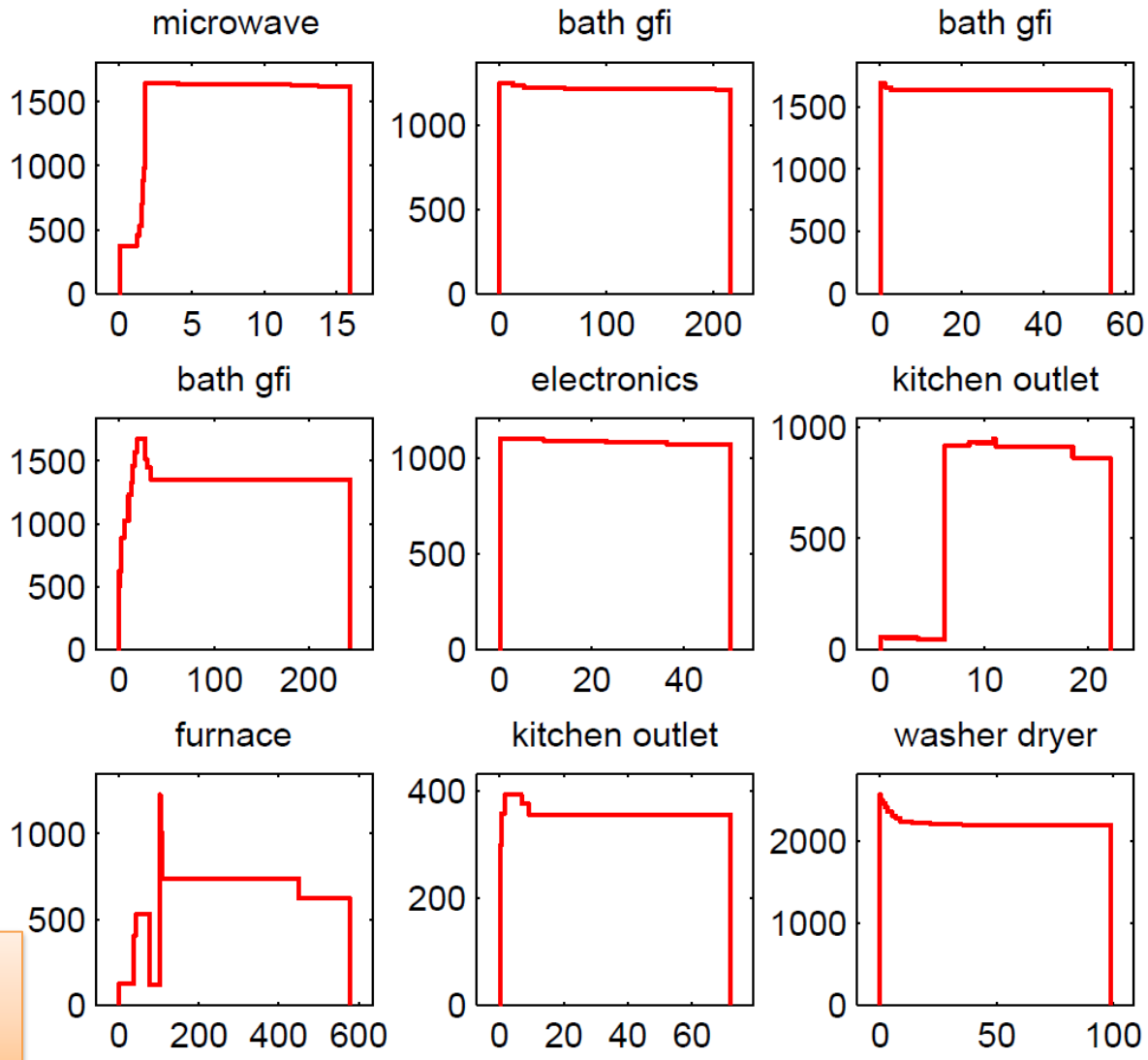


Extract all potential device snippets from signal (power coming on and going off)

Compute similarities between snippets and use spectral clustering

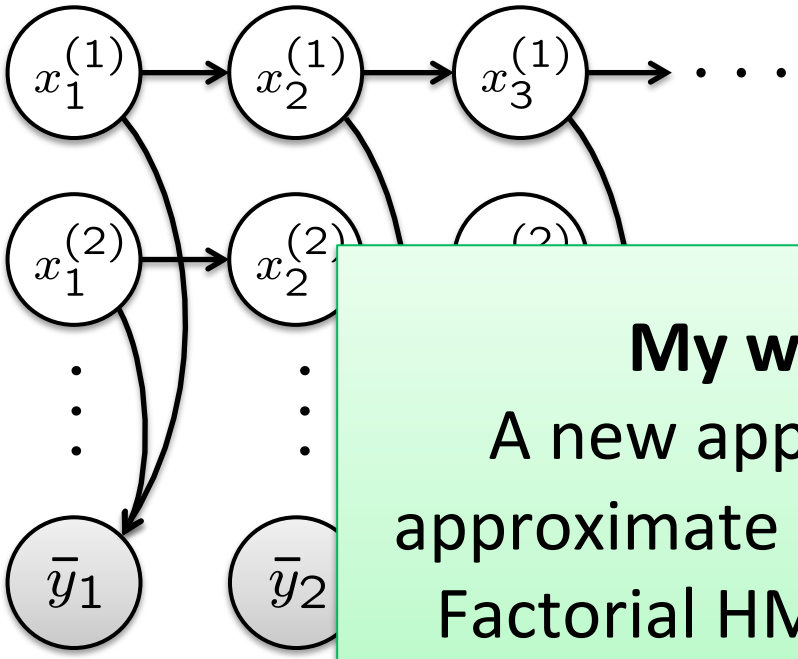
2. ***Learning*** (unsupervised)

Learning with Spectral Clustering



2. ***Learning***
(unsupervised)

Inference in Factorial HMMs



Given $\bar{y}_{1:T}$, find $x_{1:T}^{(1:N)}$

ence is intractable

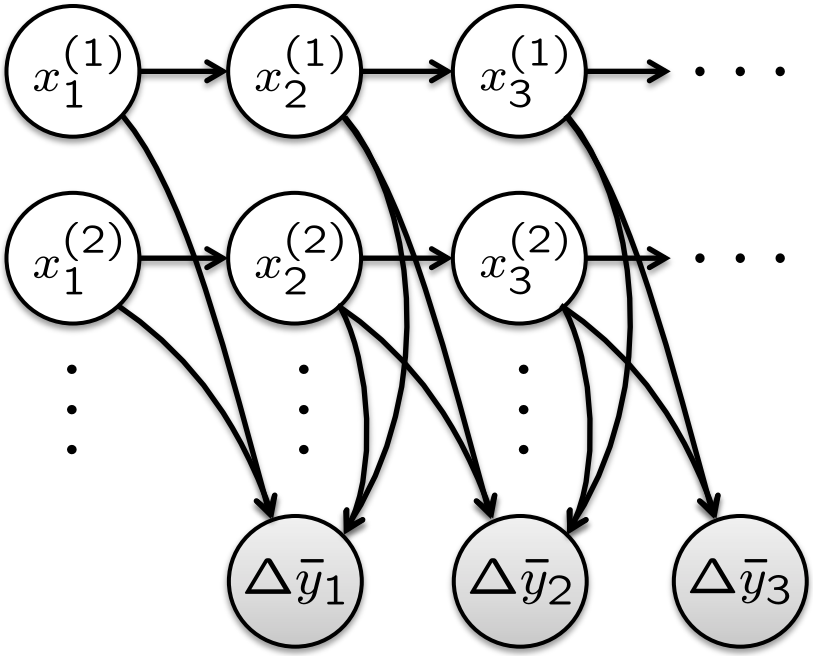
~10 states each →
otal states)

My work:
A new approach to approximate inference in Factorial HMMs based upon *convex optimization*

3. ***Inference***

Addition #1: Signal Differences

- Rather that look only at absolute signal, look at ***differences*** between successive levels



$$\Delta\bar{y}_t | x_t^{(1:N)} \sim \mathcal{N}\left(\sum_{i=1}^{N} \mu^{(i)}_{x^{(i)}_{t+1}} - \mu^{(i)}_{x^{(i)}_{t}}, \Sigma\right)$$

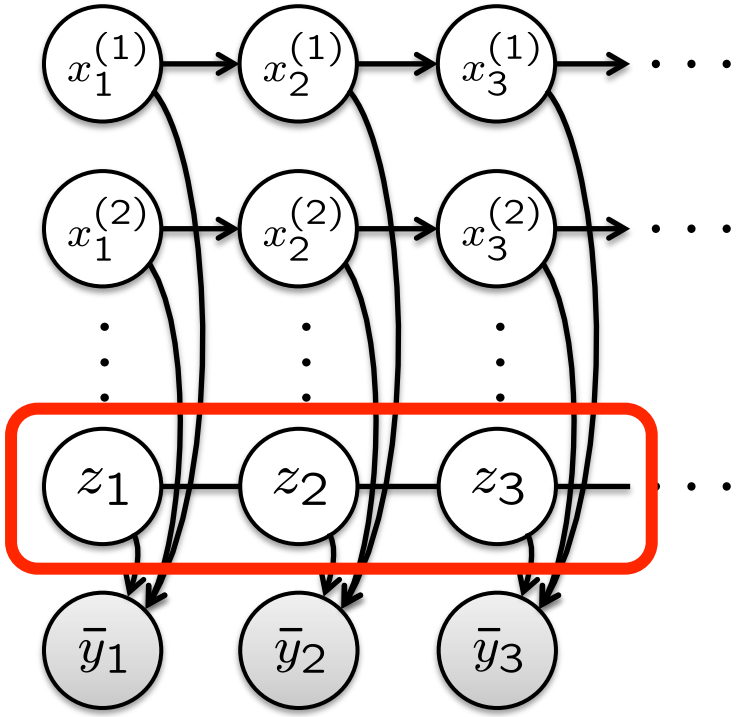
Would expect this be useful when devices change state one at a time

1. ***Specify*** model

3. ***Inference***

Addition #2: Robust HMM

- Not going to have a model for every device in the home, allow ***unassigned*** power



$$p(z_{1:T}) \propto \exp\left\{-\lambda \sum_{t=1}^{T-1} |z_{t+1} - z_t|\right\}$$

Total Variation penalty, allows for arbitrary unknown power, but favors ***piecewise-constant*** signals

1. ***Specify*** model

3. ***Inference***

Inference as Optimization

- Most likely assignment (MAP inference) can now be cast as mixed-integer QP

Additive FHMM

$$\begin{aligned} \text{maximize} \quad & \log p(x_{1:T}^{(1:N)}, z_{1:T} | \bar{y}_{1:T}) \\ \text{subject to} \quad & Q(x_{1:T}^{(1:N)}) \in \mathcal{L}(G) \\ & Q(x_{1:T}^{(1:N)}) \in \{0,1\} \end{aligned}$$

Difference FHMM

$$\begin{aligned} \text{maximize} \quad & \log p(x_{1:T}^{(1:N)}, \Delta z_{1:T} | \Delta \bar{y}_{1:T}) \\ & Q(x_{1:T}^{(1:N)}) \in \mathcal{L}(G) \\ & Q(x_{1:T}^{(1:N)}) \in \{0,1\} \end{aligned}$$

End result: convex approximate inference method, can be solved quickly for hundreds of thousands of variables

Enforce one-at-a-time in posterior

LP

Drop integer constraint

P

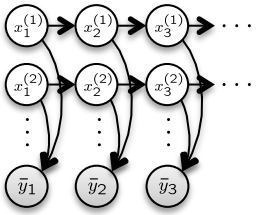
Solve Jointly

3. ***Inference***

Summary

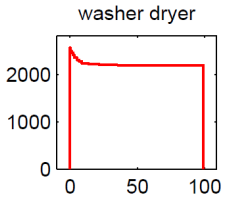
1. ***Specify*** model

Factorial HMM + 2 new additions: difference signal and robust component



2. ***Learning*** (unsupervised)

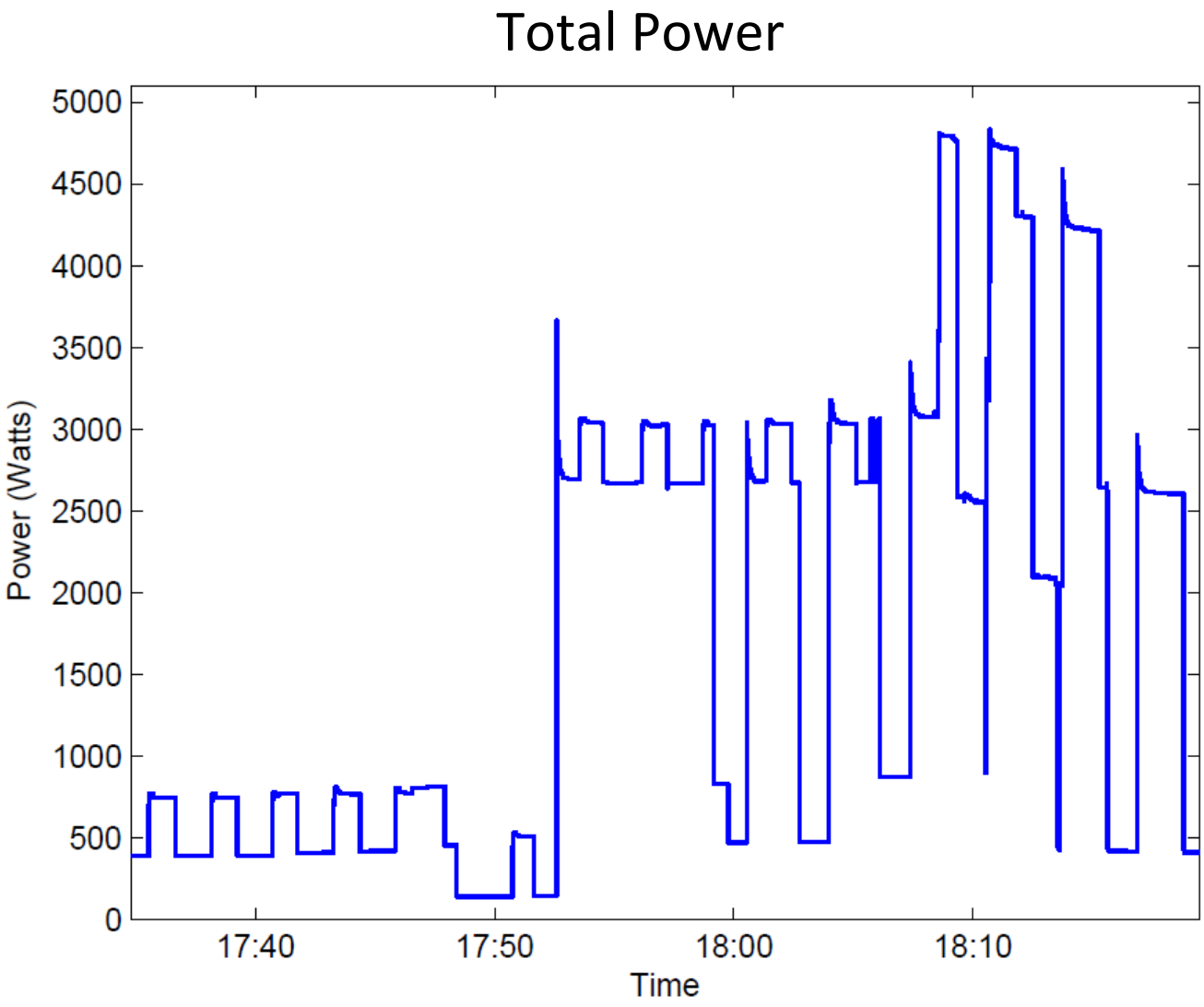
New learning approach based upon spectral clustering of device snippets



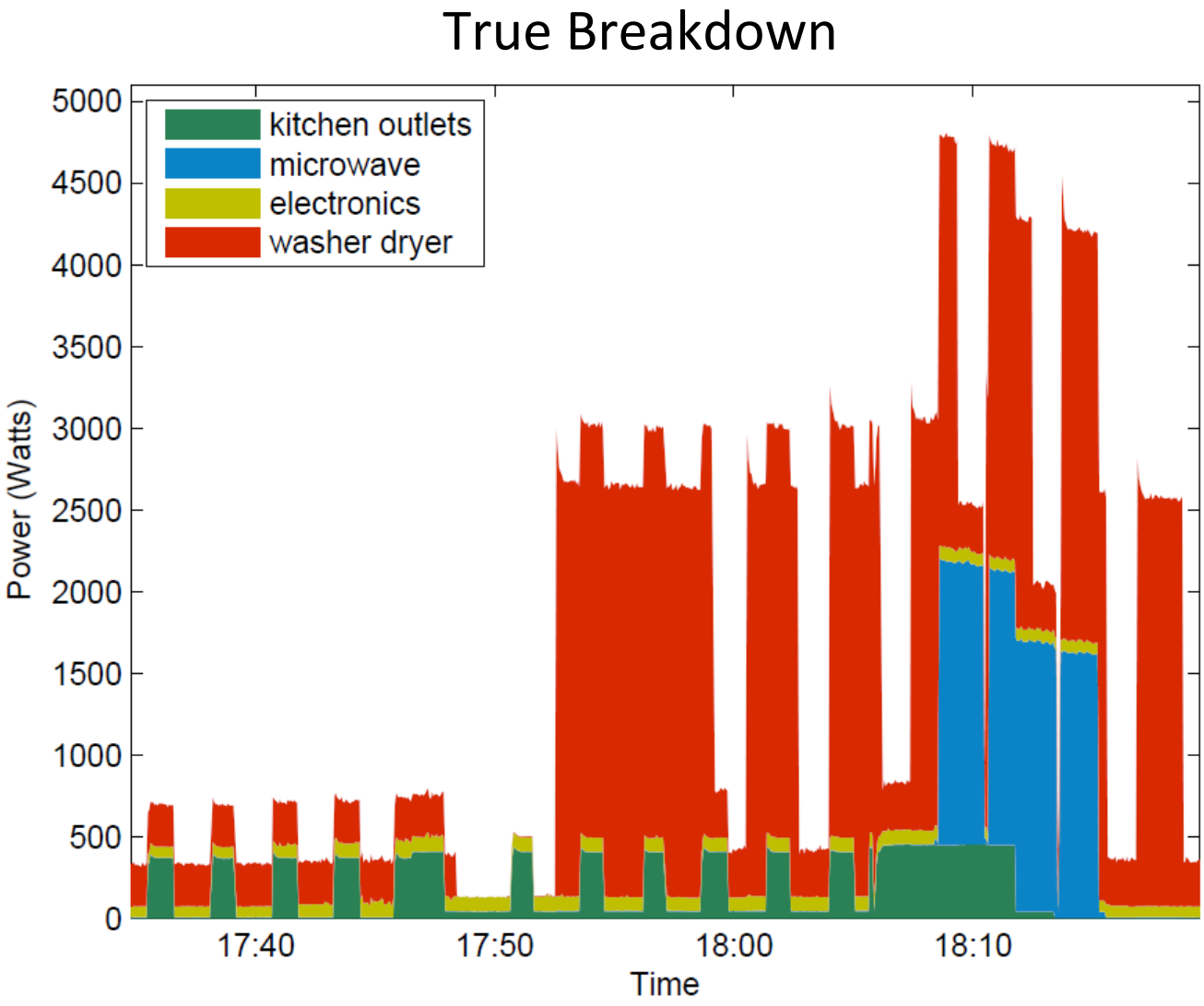
3. ***Inference***

New inference procedure based on convex optimization, relaxations

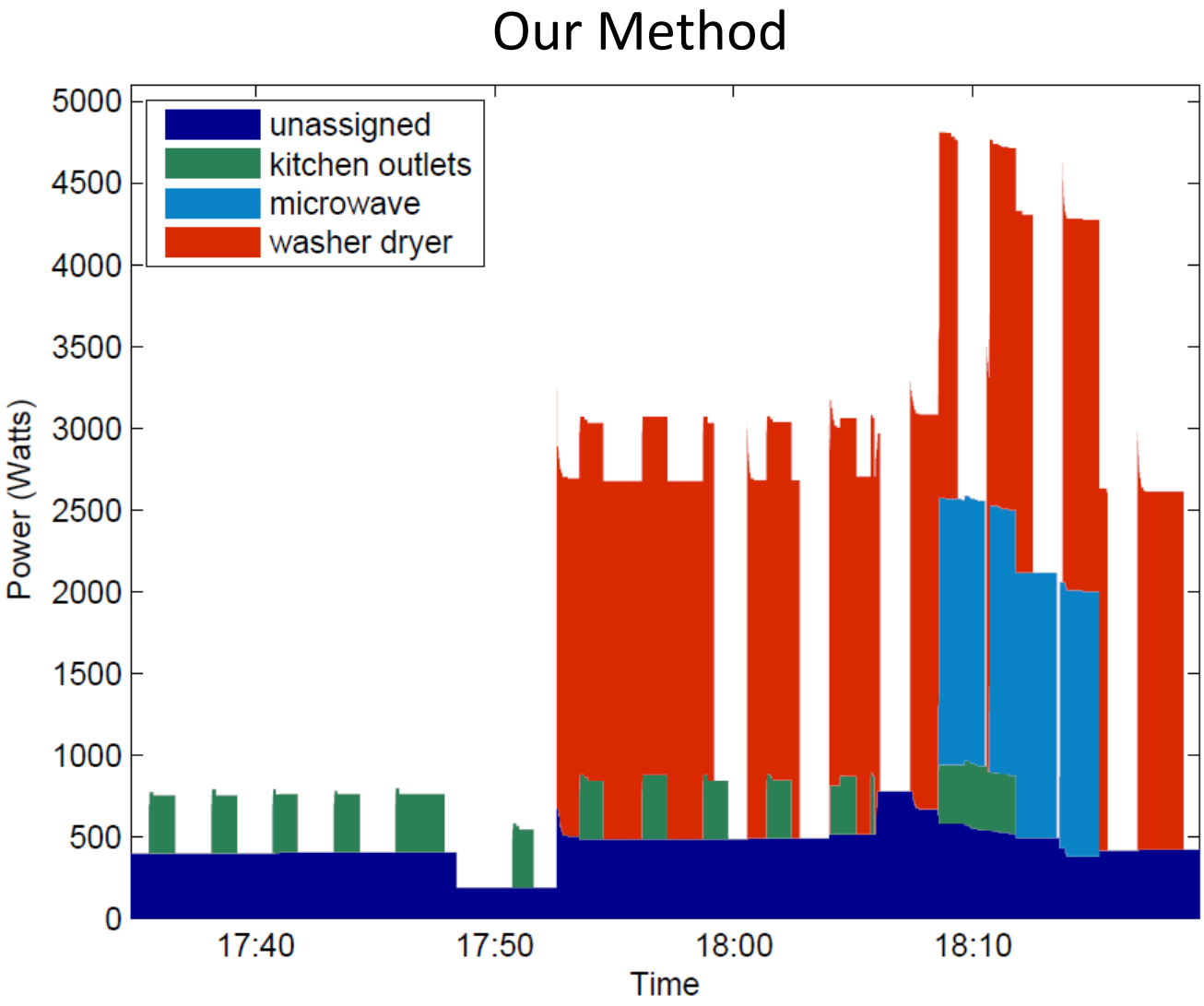
Performance



Performance

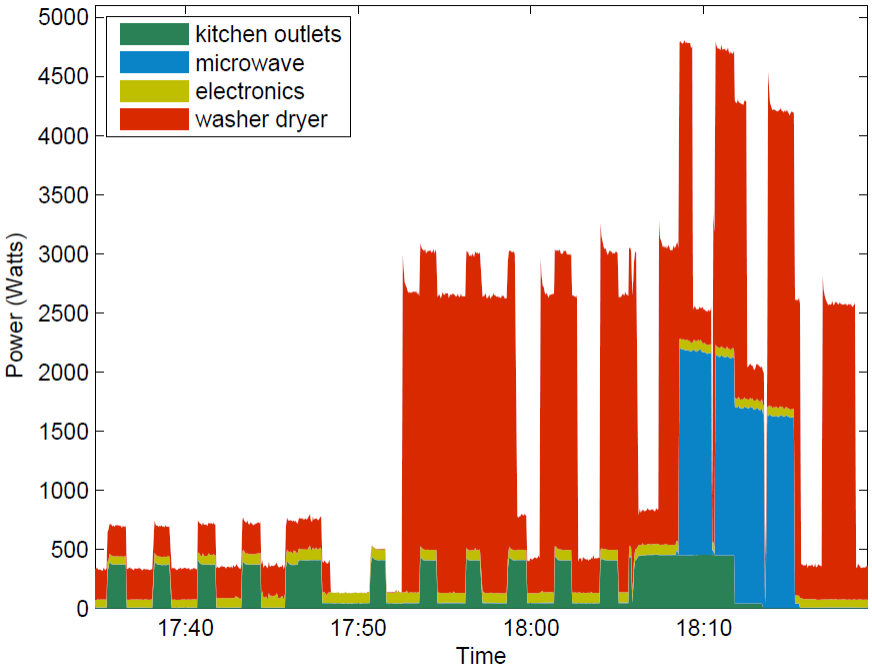


Performance

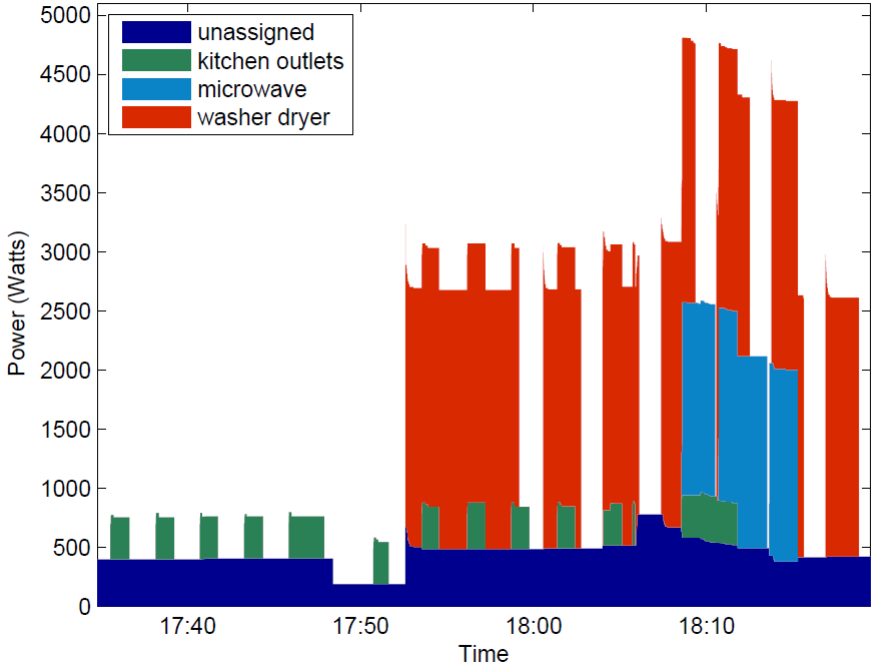


Performance

True Breakdown

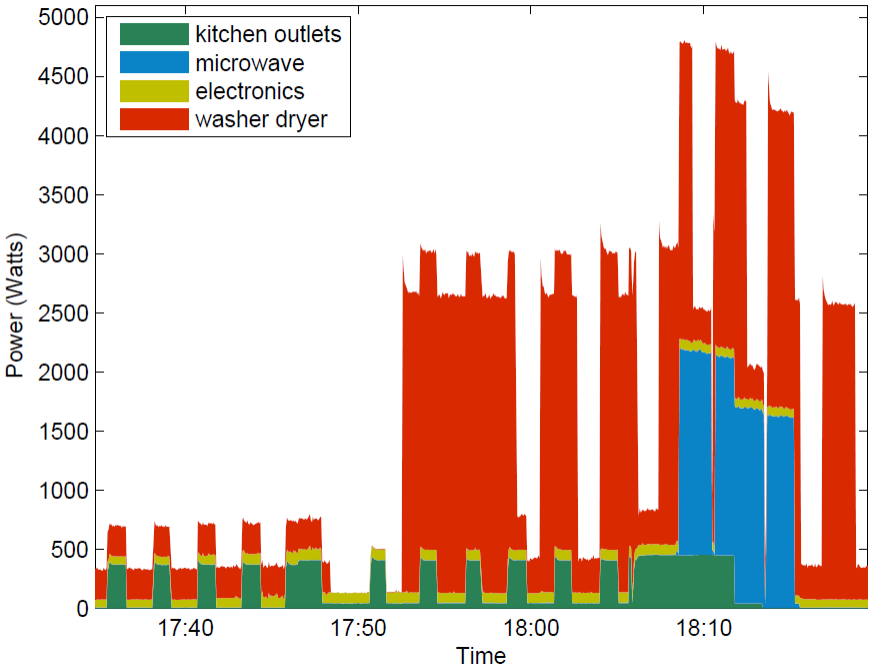


Our Method

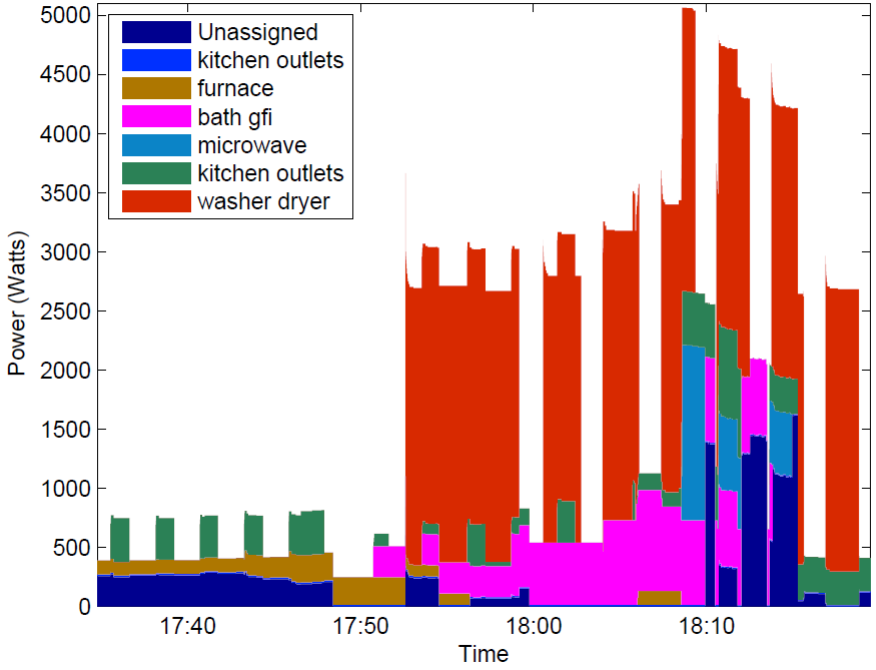


Performance

True Breakdown

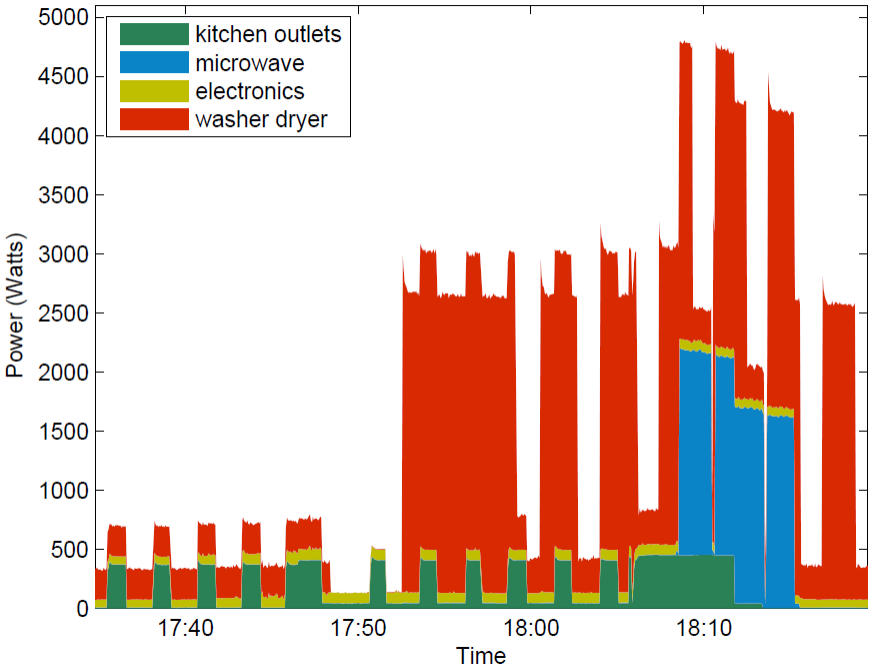


Event Based

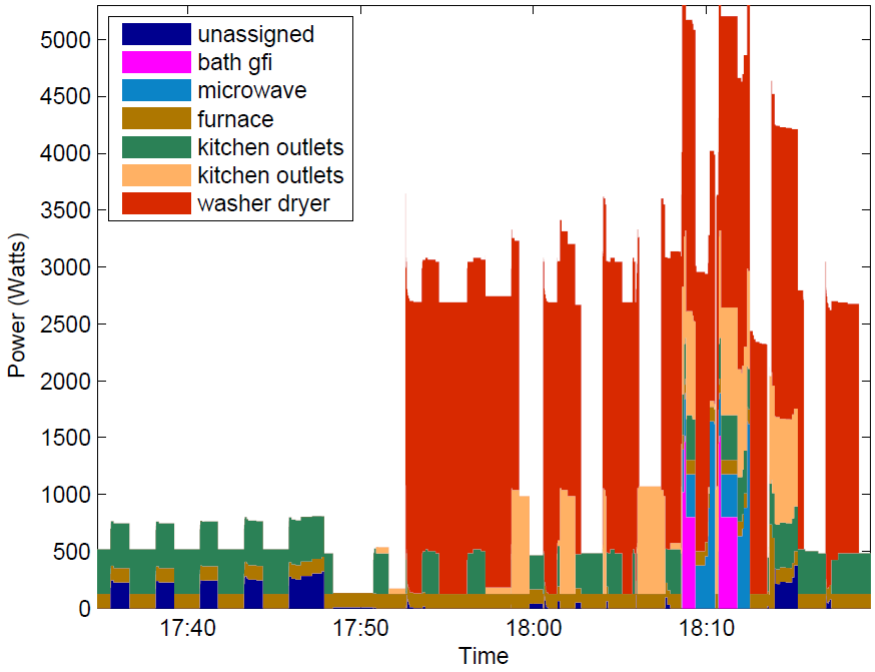


Performance

True Breakdown



Structured Mean Field



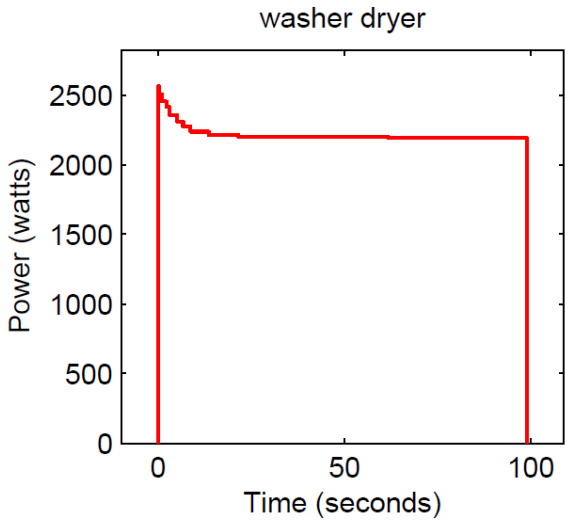
Performance Over Two Weeks

Circuit	Our Method	Structured Mean Field	Event-based
Microwave	98% / 66%	97% / 4%	98% / 28%
Bath GFI	83% / 71%	50% / 9%	23% / 21%
Kitchen Outlets	38% / 13%	10% / 48%	57% / 15%
Furnace	92% / 71%	13% / 15%	25% / 71%
Kitchen Outlets	45% / 16%	13% / 24%	27% / 11%
Washer / Dryer	99% / 73%	89% / 77%	95% / 64%
Total	**87% / 60%**	**36% / 45%**	**49% / 53%**

Performance measured by per-circuit precision/recall

Performance Over Two Weeks

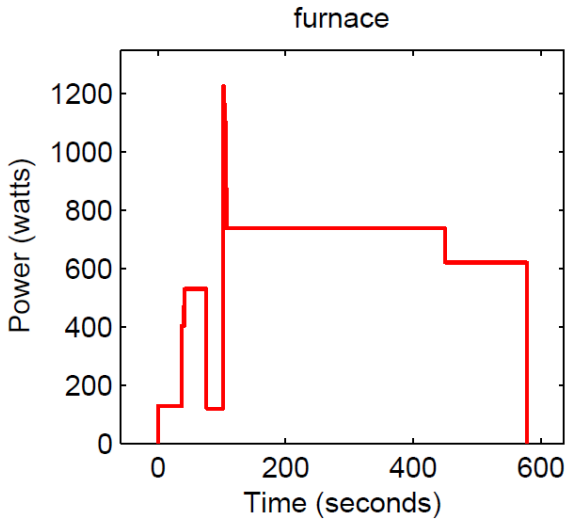
Circuit	Our Method	Structured Mean Field	Event-based
Microwave	98% / 66%	97% / 4%	98% / 28%
Bath GFI	83% / 71%	50% / 9%	23% / 21%
Kitchen Outlets	38% / 13%	10% / 48%	57% / 15%
Furnace	92% / 71%	13% / 15%	25% / 71%
Kitchen Outlets	45% / 16%	13% / 24%	27% / 11%
Washer / Dryer	99% / 73%	89% / 77%	95% / 64%
Total	**87% / 60%**	**36% / 45%**	**49% / 53%**



Performance measured by per-circuit precision/recall

Performance Over Two Weeks

Circuit	Our Method	Structured Mean Field	Event-based
Microwave	98% / 66%	97% / 4%	98% / 28%
Bath GFI	83% / 71%	50% / 9%	23% / 21%
Kitchen Outlets	38% / 13%	10% / 48%	57% / 15%
Furnace	92% / 71%	13% / 15%	25% / 71%
Kitchen Outlets	45% / 16%	13% / 24%	27% / 11%
Washer / Dryer	99% / 73%	89% / 77%	95% / 64%
Total	**87% / 60%**	**36% / 45%**	**49% / 53%**



Performance measured by per-circuit precision/recall

Directions for Future Work

- Iterated learning and inference, using initial models as priors
- Integrate with supervised learning procedures to automaically classify devices
- Integrate with state-of-the-art sensing and feature extraction
- Testing on additional homes (expanding REDD to 50 homes: http://redd.csail.mit.edu)