How well can HMM model load signals

3rd International Workshop on Non-Intrusive Load Monitoring, May 14th, Vancouver, Canada

Lukas Mauch, Karim Said Barsim and Bin Yang

Institute of Signal Processing and System Theory
University of Stuttgart

Content

How to model loads

Hidden Markov Models (HMM)

- Basics
- HMM states vs. load states
- Restricting the HMM parameters

Model selection

- Akaike Information Criterion (AIC)

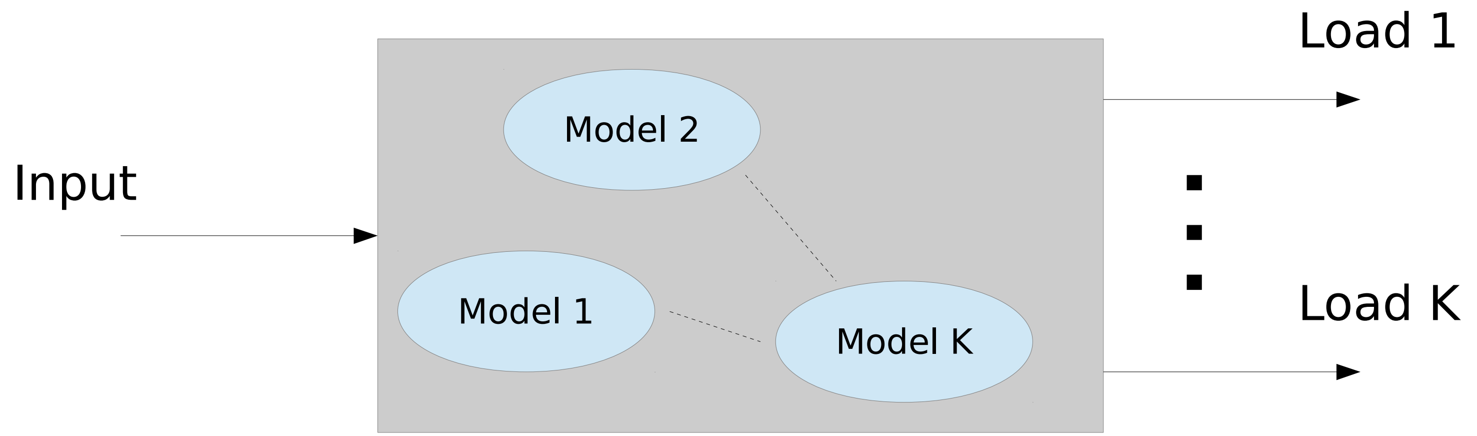
Model adaptation

- An easy parametric transformation

Experiments

How to model loads

Non-Intrusive load monitoring as a single channel source separation problem



Data acquisition

- Single channel
- Low frequency
- Real power only
- Hierarchical time dependencies
- Non-stationary

Separation

- Factorization/clustering methods
- Methods for denoising
- Bayesian methods

Main problem

- We need suited signal models to perform separation

How to model loads

State of the art

Piecewise modelling

- Event based approaches
- Problem of segmentation
- Loss of information for variable loads
- Captures little information about time dependencies

Recurrent Neural Network

- Very powerful model
- Can learn hierarchical time dependencies
- Hard to train
- Hard to interprete

Hidden Markov Model

- Simple model that can capture dependencies between adjacent states
- Easy to train
- Good results if used with fHMM
- No investigation yet how well they fit to load signals
- Good to interprete?

Questions related to HMM

Goodness of fit

- Are HMMs suited to model all kind of loads?
- How can we interprete the HMM states? Are they equal to the physical load states?
- How to choose the number of states?

Model adaptation

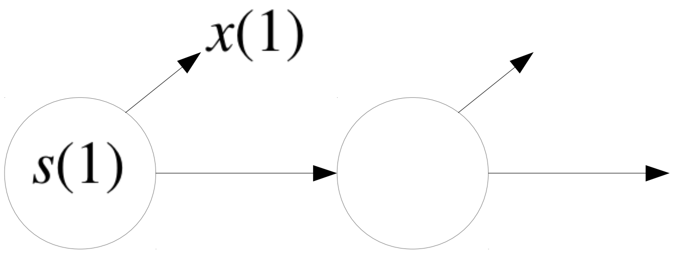
- Can we adapt HMMs to other houses?

Hidden Markov Models

Basics

Observation sequences and state sequences

$$X = \{x(1), \ldots, x(N)\} \qquad x(n) \in \mathbb{R}$$
$$S = \{s(1), \ldots, s(N)\} \qquad s(n) \in \{1, \ldots, M\}$$



Initial state and state transition probabilities

$$\pi_i = P(s(1) = i), \; 1 \leq i \leq M, \; \underline{\pi} = [\pi_1, \ldots, \pi_M]^T \in \mathbb{R}^M$$

$$a_{ij} = P(s(n) = i | s(n-1) = j), \; 1 \leq i, j \leq M$$
$$\mathbf{A} = [a_{ij}] \in \mathbb{R}^{M \times M}$$

Emission probabilities

$$\rho_i(x(n)) = p(x(n) | s(n) = i), \; 1 \leq i \leq M$$

$$\rho_i(x(n)) \sim N(\mu_i, \sigma_i^2)$$

Hidden Markov Models

Basics

Joint sequence probability

$$p(X, S\,|\underline{\theta}) = \pi_{s(1)}\rho_{s(1)}(x(1)) \prod_{n=2}^{N} a_{s(n)s(n-1)}\rho_{s(n)}(x(n))$$

Training and state inference

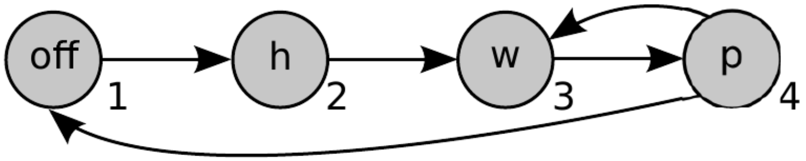
$$\hat{\underline{\theta}} = \arg\max_{\underline{\theta}} \prod_{k} p(X_k|\underline{\theta})$$ Baum-Welch algorithm

$$\hat{S} = \arg\max_{S} p(X, S\,|\hat{\underline{\theta}})$$ Viterbi algorithm

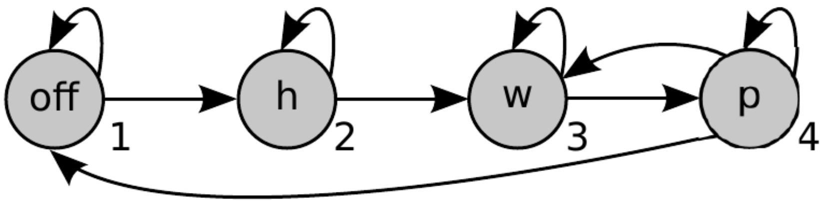
Hidden Markov Models

HMM states vs load states

Transition of load states



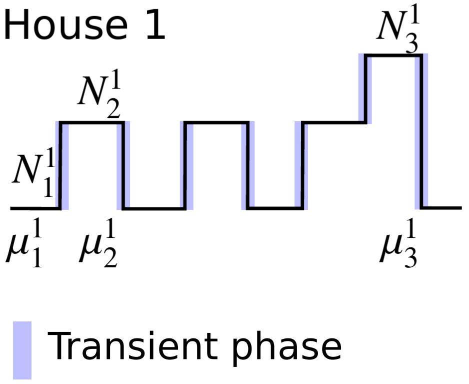
Transition of HMM states



- In general the load states and HMM states are different
- Their relationship depends on
  - Type of the load
  - Sampling frequency
  - Transient phase of the load

- HMM states = load states if
  - States with perfectly constant power consumption
  - Sharp transient phase

House 1



Hidden Markov Models

Controlled vs uncontrolled

controlled multi-state loads	uncontrolled multi-state loads	variable loads
washing machine dishwasher microwave toaster	fridge lighting ciruit	PC amplifier

Restricting the HMM parameters

- For some controlled loads we can use prior knowledge to reduce the number of HMM parameters
  → better estimate of parameters

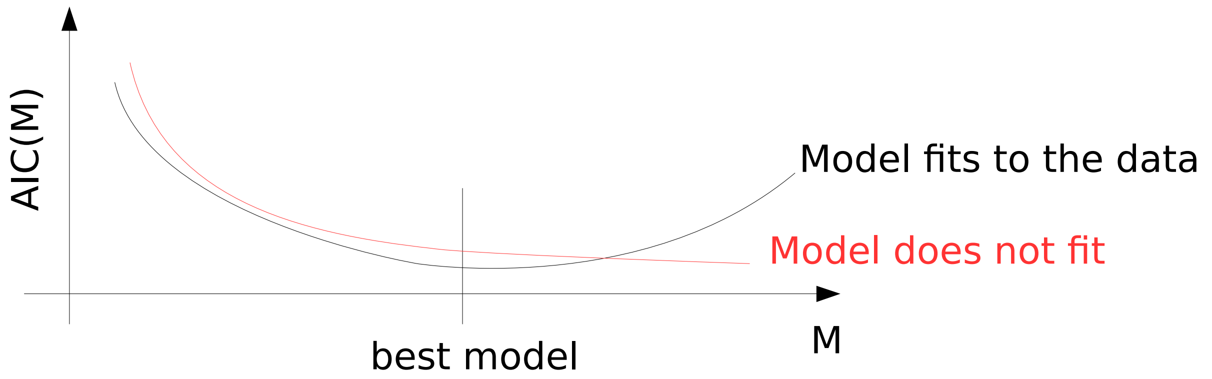
- Example: periodic chain structure leads to special transition matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 & a_{14} \\ a_{21} & a_{22} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{pmatrix}, \quad a_{ij} \neq 0.$$

Model selection

Akaike Information Criterion (AIC)

$$AIC(M) = -2L_M + 2N_M, \;\; L_M = \ln p(X|\underline{\hat{\theta}}_M)$$

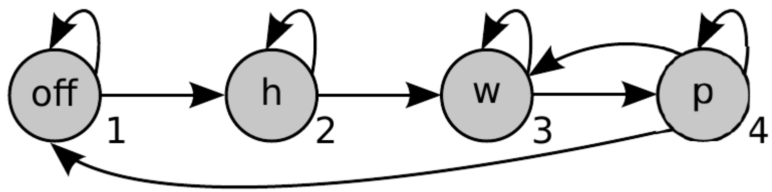


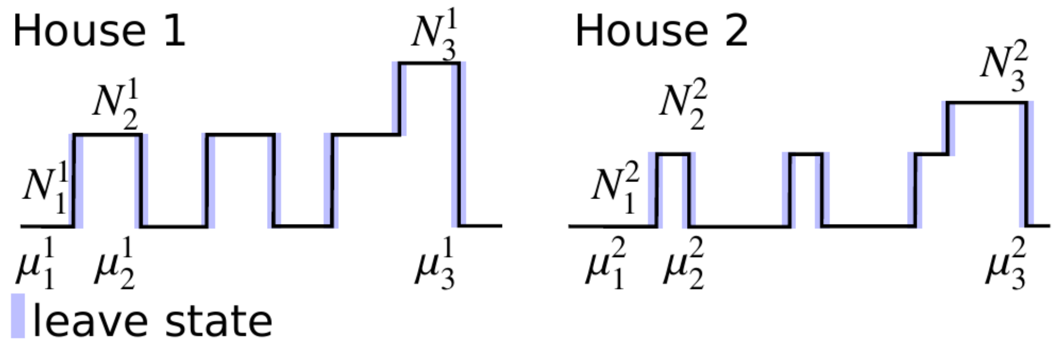
- Measure how well a model fits to a specific load
- Balances goodness of fit (data likelihood) against model complexity (number of parameters M)
- Choose model with lowest AIC

- If model does not fit
  - Increasing model complexity always leads to increasing data likelihood

Model adaptation

Basics





What are causes for differences between signals of loads of the same kind

- Differences in sampling frequency
- Differenent power consumption in each state
- Different state duration

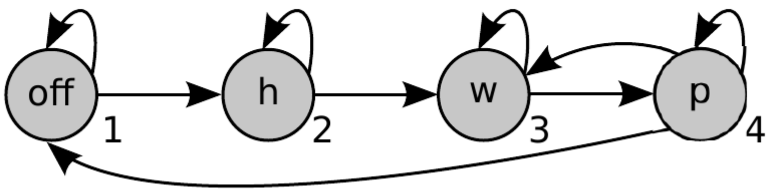
Assumptions

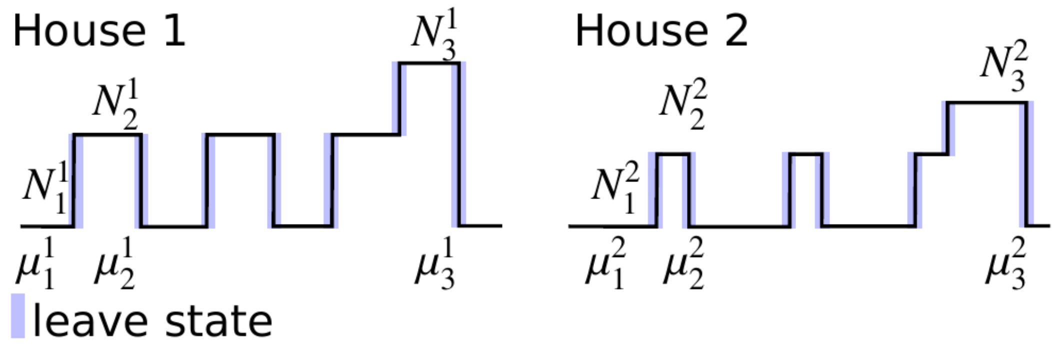
- Periodic signal patterns
- Loads of the same kind share the same set of states
  → We can use a simple transformation of parameters to adapt the HMM

$$\underline{\hat{\theta}}_2 = f(\underline{\hat{\theta}}_1, \underline{\alpha})$$

Model adaptation

Adaptation of the state mean





What are causes for differences between signals of loads of the same kind

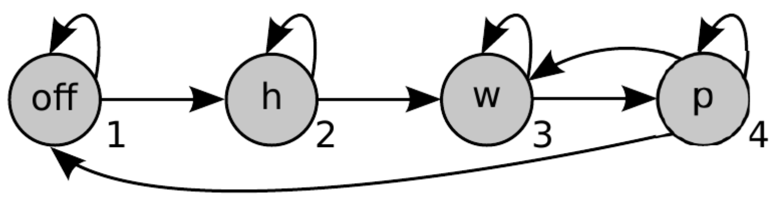
- Differences in sampling frequency
- Differenent power consumption in each state
- Different state duration

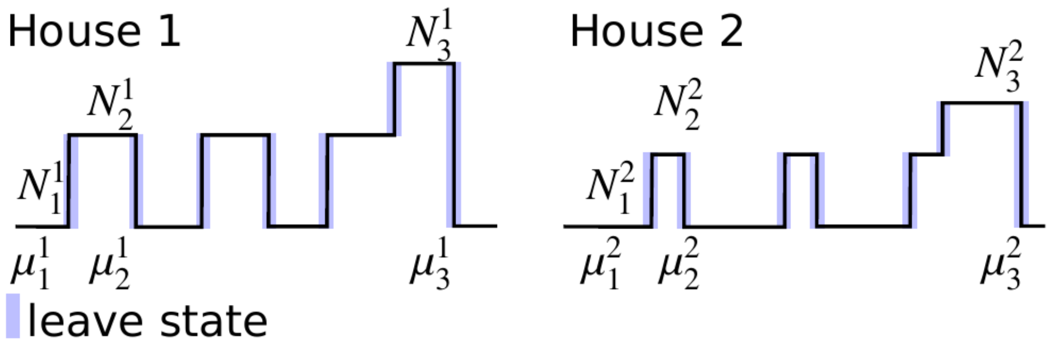
→ The means of each emission probability of all states are scaled independently

$$\mu_i^2 = \alpha_i \mu_i^1, \quad 1 \le i \le M.$$

Model adaptation

Adaptation of the transition matrix





What are causes for differences between signals of loads of the same kind

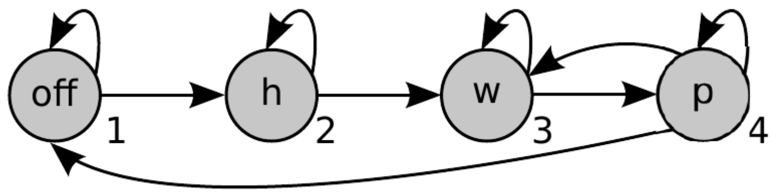
- Differences in sampling frequency
- Differenent power consumption in each state
- Different state duration

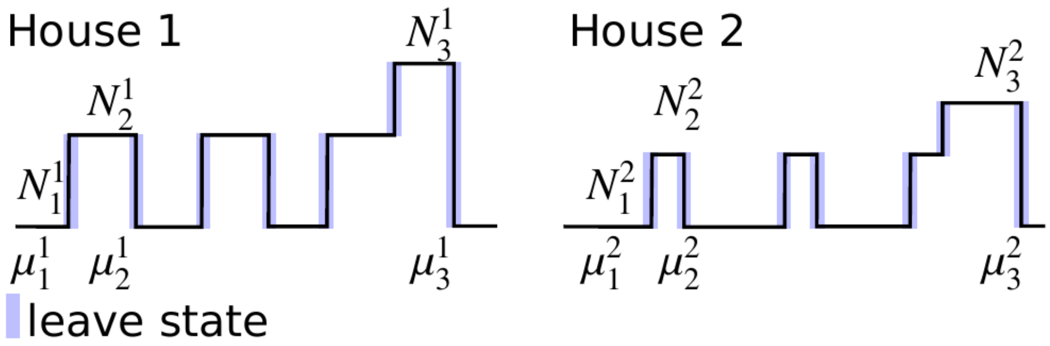
$$N_i^1 \rightarrow N_i^2 = c_i N_i^1$$
$$\mathbf{A}^1 \rightarrow \mathbf{A}^2$$

$$
\begin{aligned}
a_{ii}^k &= P^k(s(n) = i|s(n-1) = i) = \frac{N_i^k - 1}{N_i^k},\ 1 \leq i \leq M \\
a_{ij}^k &= P^k(s(n) = i|s(n-1) = j) \\
&= P^k(s(n) = i|l_j(n))\, \underbrace{P^k(s(n) \neq j|s(n-1) = j)}_{1-a_{jj}^k},\ i \neq j
\end{aligned}
$$

Model adaptation

Adaptation of the transition matrix





What are causes for differences between signals of loads of the same kind

- Differences in sampling frequency
- Differenent power consumption in each state
- Different state duration

$$N_i^1 \rightarrow N_i^2 = c_i N_i^1$$
$$\mathbf{A}^1 \rightarrow \mathbf{A}^2$$

Re-scaling the diagonal elements

$$a_{ii}^2 = \frac{N_i^2 - 1}{N_i^2} = \underbrace{\frac{c_i N_i^1 - 1}{c_i (N_i^1 - 1)}}_{\lambda_i} a_{ii}^1, \quad 1 \leq i \leq M$$

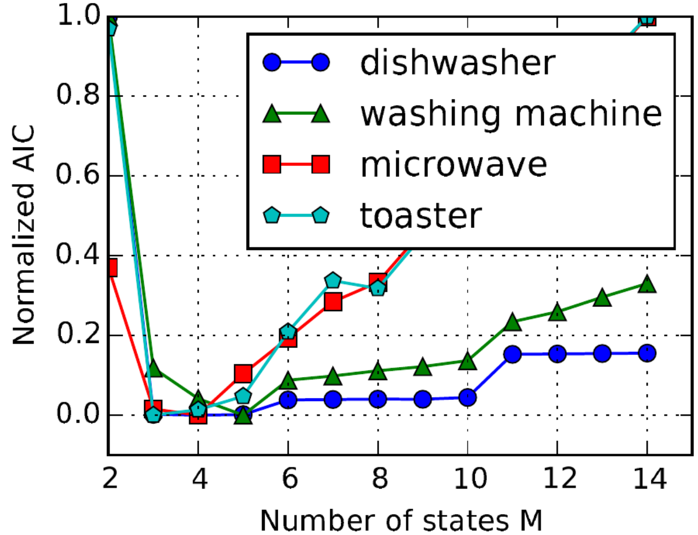
Re-scaling the off-diagonal elements

$$\begin{aligned} a_{ij}^1 &= P^2(s(n) = i | l_j(n))(1 - a_{jj}^2) \\ &= P^1(s(n) = i | l_j(n))(1 - a_{jj}^2)\frac{1 - a_{jj}^2}{1 - a_{jj}^1} \\ &= \frac{1 - \lambda_i a_{11}^1}{1 - a_{ii}^1}, \quad i \neq j \end{aligned}$$

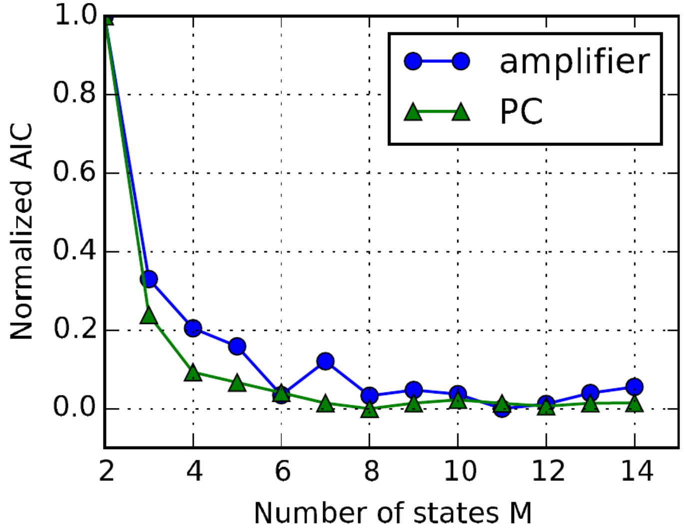
Experiments

AIC for different loads

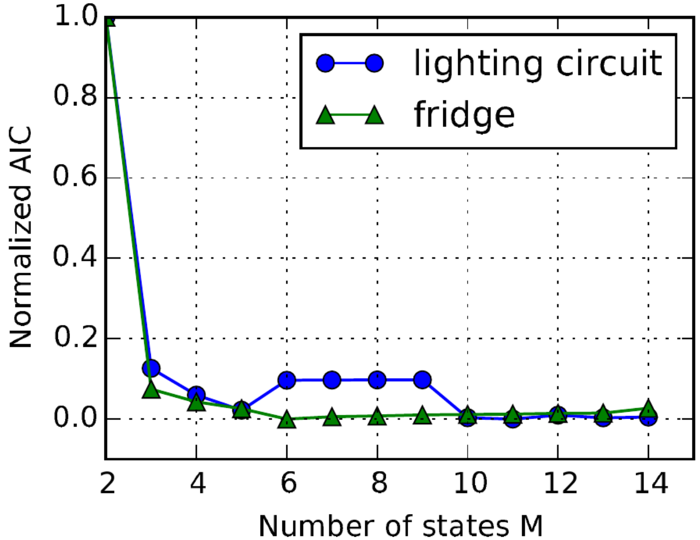
Controlled multi-state load



Variable load



Uncontrolled multi-state load

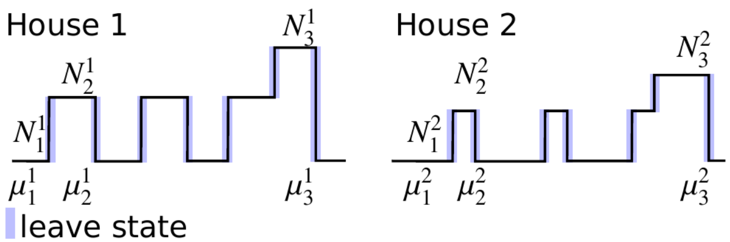


Result

- Clear minima for controlled multi-state load using only few HMM states
- AIC keeps decreasing for increasing number of states (increasing model complexity) in case of uncontrolled multi-state and variable loads

Experiments

Adaptation to simulated data



	μ_1^2	μ_2^2	μ_3^2	a_{11}^2	a_{22}^2	a_{33}^2
true	0	1.2	2.3	0.875	0.95	0.955
estimated	0	1.19	2.3	0.875	0.94	0.97
	μ_1^1	μ_2^1	μ_3^1	a_{11}^1	a_{22}^1	a_{33}^1
orig.	0	1	2	0.8	0.9	0.95

- Simulated periodic load signal
- Train on 350 samples of house 1 (10 periods)
- Adapt on 35 samples of house 2 (1 period)

Adaptation to measured data

load	$\ln p(X_1 \vert \underline{\hat{\theta}}_1)$	$\ln p(X_2 \vert \underline{\hat{\theta}}_1)$	$\ln p(X_2 \vert f(\underline{\hat{\theta}}_1, \underline{\alpha}))$
fridge	3311	1430	3005
dishwasher	2881	-1614	2453
washing machine	7100	2369	3375

Conclusion

Is HMM a suited model for all loads?

- Good model for controlled multi-state loads with fixed periodic behaviour
- Bad model for uncontrolled multi-state and variable loads

Can we adapt HMM to different houses?

- For periodic signals that share the same set of states between houses
  → parametric transformation of model parameters
  can be used for adaptation
- Only little data for adaptation is needed

Outlook

- For which loads can we reduce the model complexity by restricting the model parameters?
- How do we have to modify the HMM assumptions to get good models for variable and uncontrolled multi-state loads?