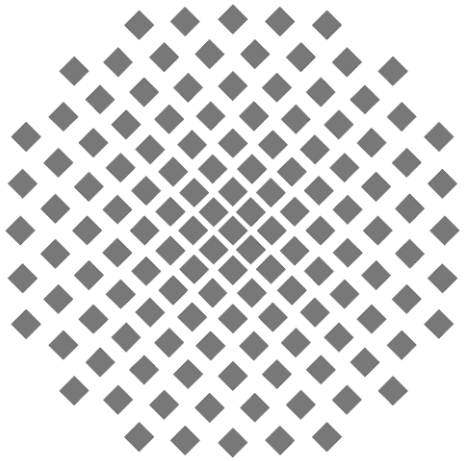
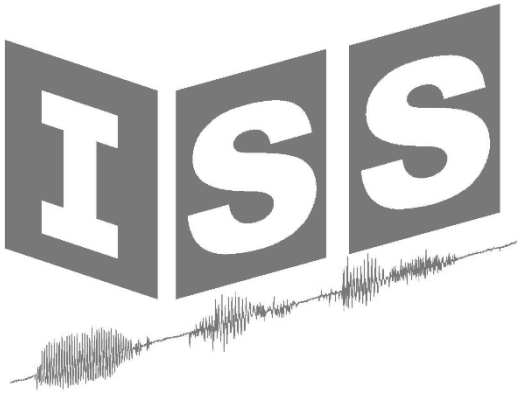
Neural Network Ensembles to Real-time Identification of Plug -level Appliance Measurements





NILM Workshop 2016, 14th May

Karim Said Barsim, Lukas Mauch, and Bin Yang

Institute of Signal Processing and System Theory

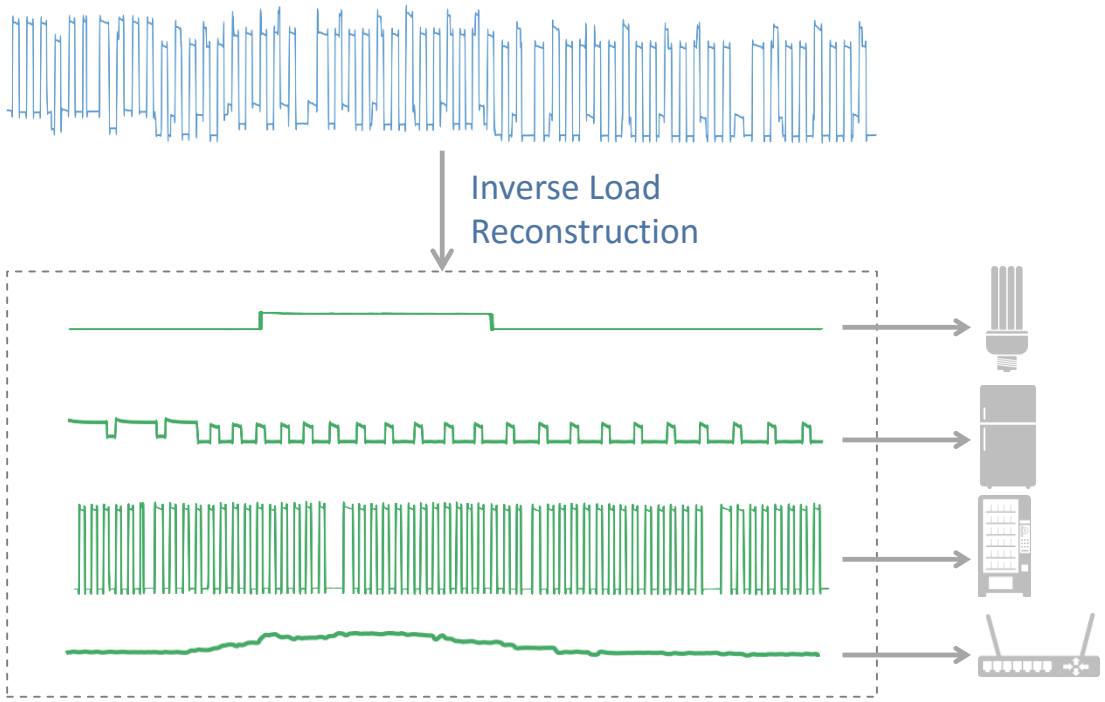
University of Stuttgart, Stuttgart, Germany

{karim.barsim,lukas.mauch,bin.yang}@iss.uni-Stuttgart.de

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Introduction and motivation

- Appliance identification
  - A sub-task of the NILM problem
  - Sometimes a disjoint process
- An upper bound
  - Raw signatures (current $i(n)$ and voltage $v(n)$)
  - High resolution signals
  - Plug-level measurements
- Plug-level appliance identification
  - Smart outlets
- Aspects of the identification process
  - Training data
  - Sampling frequency
  - Signature variations



L. Mauch
Stuttgart University

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Task and tools

Generic appliance identification from high resolution raw measurements

- Neural network ensembles:
  - Suitable for raw-data learning
  - Unstable w.r.t training data (i.e. suited for models ensembles)

- Plug Load Appliance Identification Dataset (PLAID) [1]
  - Plug-level raw measurements
  - Generic appliance categories

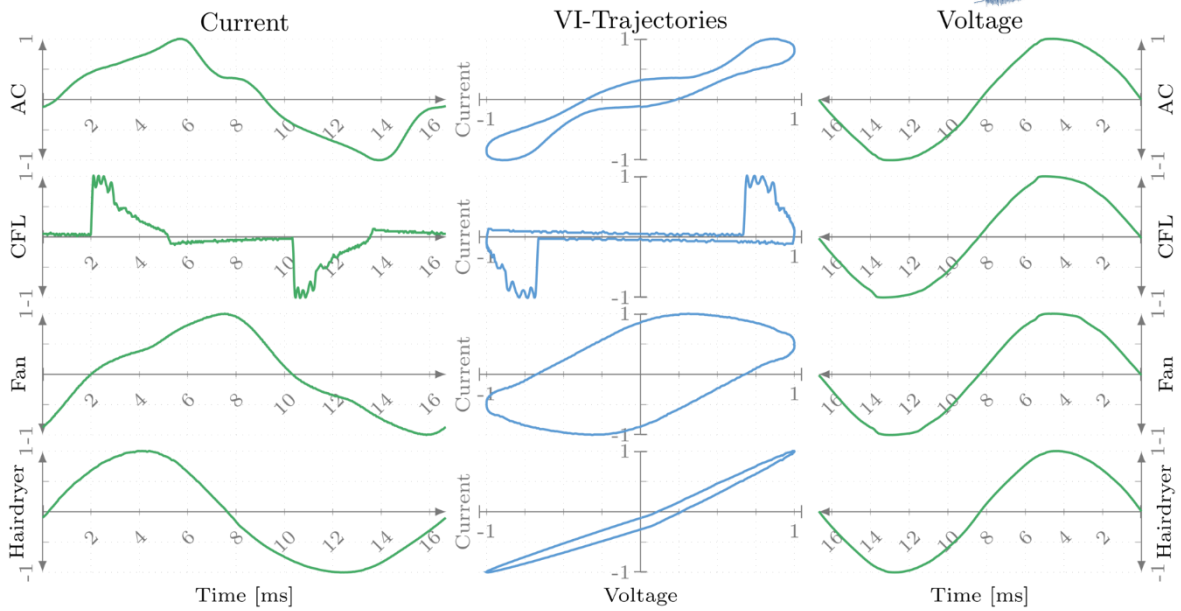
L. Mauch
Stuttgart University

[1] J. Gao, S. Giri, E. C. Kara, and M. Berg´es, “PLAID: A Public Dataset of High-resolution Electrical Appliance Measurements for Load Identification Research: Demo Abstract”. In proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings:198–199, 2014.

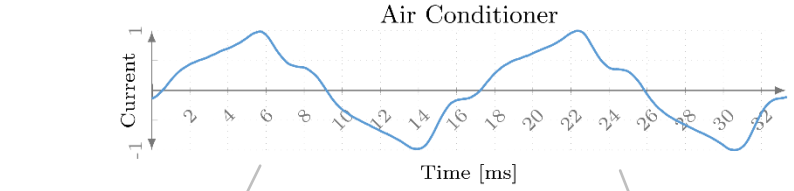
Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Appliances' Signature

- Raw current $i(n)$ and voltage $v(n)$
  - Unsupervised feature learning
  - $\boldsymbol{x}_\tau = [\hat{\boldsymbol{i}}_\tau^T,\ \hat{\boldsymbol{v}}_\tau^T]^T$
- Single period (~ 0.017 seconds @ 60Hz)
  - Real-time identification
  - $\boldsymbol{i}_\tau = [i(\tau), i(\tau+1), \ldots, i(\tau+d+1)]^T$
- Segment-based normalization
  - Discarding amplitude information
  - Generic labeling
  - $\hat{\imath}_\tau = \frac{2\,\boldsymbol{i}_\tau - (\max \boldsymbol{i}_\tau - \min \boldsymbol{i}_\tau)\,\mathbf{1}_{d,1}}{(\max \boldsymbol{i}_\tau - \min \boldsymbol{i}_\tau)}$
- Algorithmic expansion
  - Multiple phase shifts from 2 periods
  - Translation invariance and robustness to variations
  - $\mathcal{X} = \left\{ \boldsymbol{x}_\tau \,\middle|\, \tau = \tau_o + m\,\epsilon\,,\quad 0 \le m < \frac{d}{\epsilon} \right\}$



Appliance signature (green curves)



Expansion of training data

L. Mauch
Stuttgart University

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Prediction model

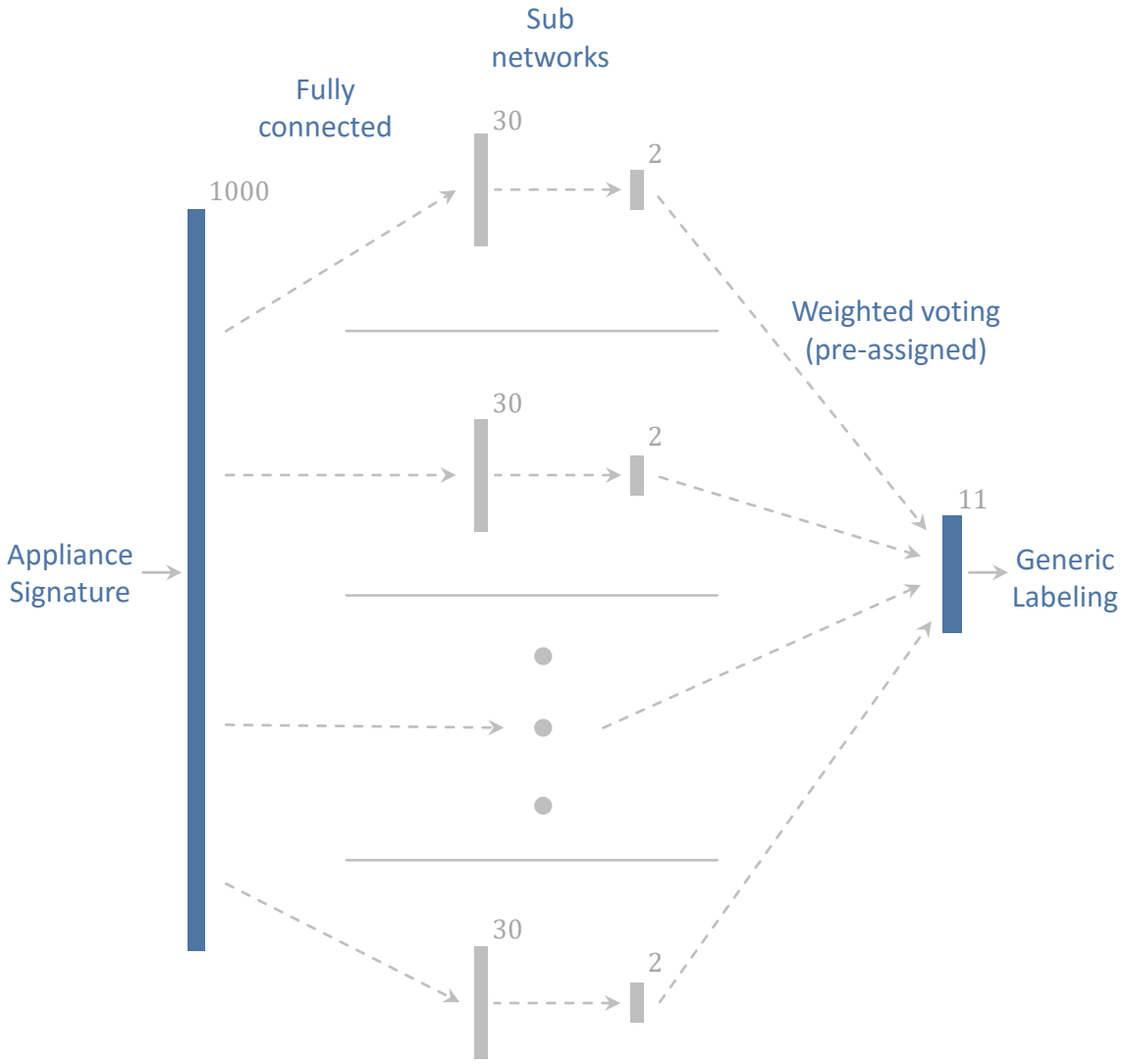
- A neural network ensemble
  - Unstable models
  - Similar to Bootstrap aggregation
  - Ex. *binary classification networks*
  - Ex. *per class combination* $\binom{11}{2} = 55$ nets
  - Ex. *shallow, feedforward, fully connected nets*

$$D_{\omega_i,\omega_j} = \{ \boldsymbol{x} \in \mathcal{X} | \omega(\boldsymbol{x}) = \omega_i \ \text{or} \omega(\boldsymbol{x}) = \omega_j \}$$

$$\hat{\theta}_{\omega_i,\omega_j}(\boldsymbol{x}) = \left[\hat{p}_{\omega_i,\omega_j}, \quad \hat{p}_{\omega_j,\omega_i} \right]$$

- Confidence-weighted voting

$$\hat{\omega}(\boldsymbol{x}) = \arg\max_{\omega_i \in \Omega} \sum_{j \neq i} \hat{p}_{\omega_j,\omega_i}(\boldsymbol{x})$$



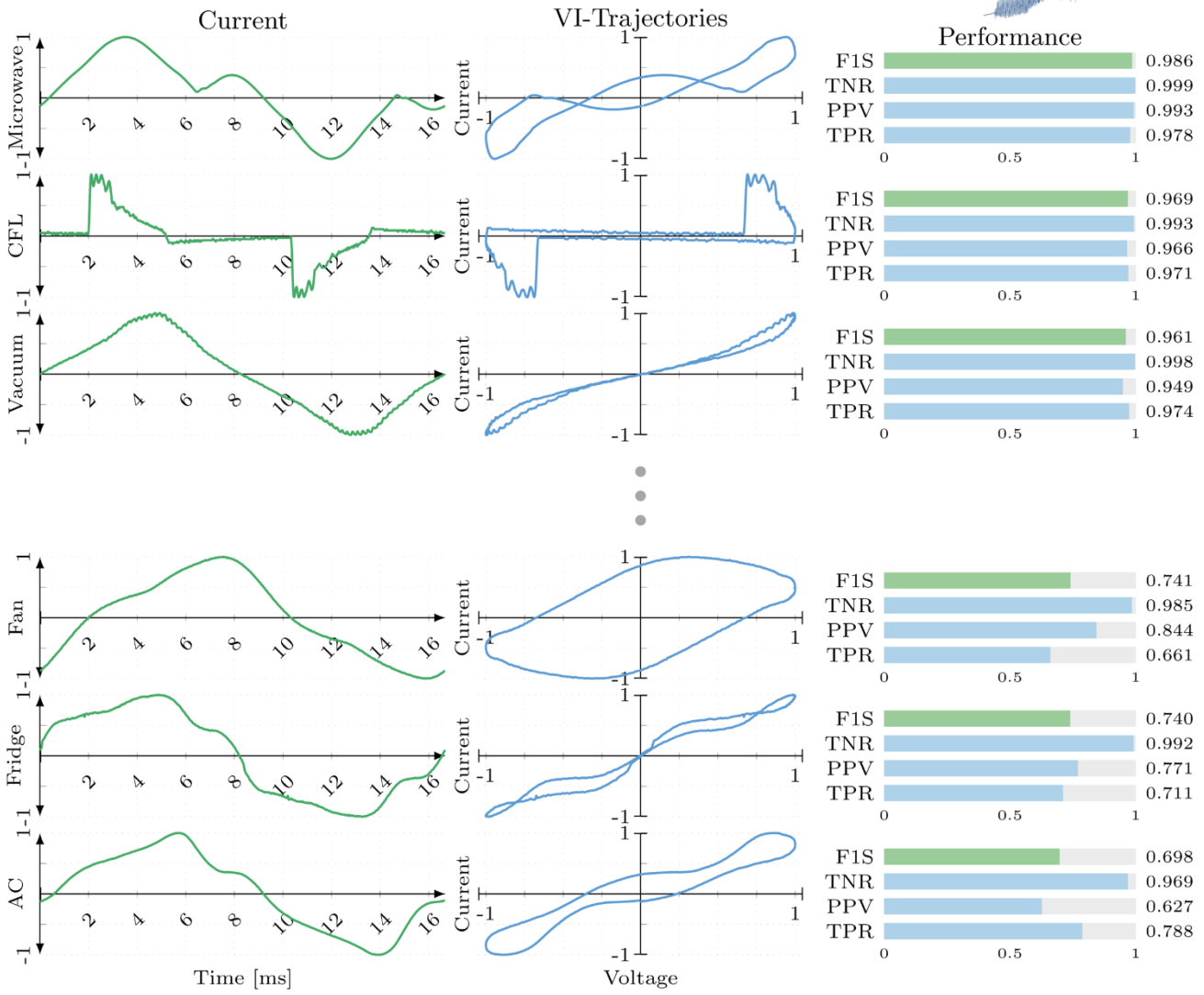
An example of the adopted prediction model

L. Mauch
Stuttgart University

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Experiments (1) - Prediction

- PLAID Dataset
  - +200 instances
  - +1000 measurements
  - Category-based (11 cat.)
  - High frequency (30 kHz)
  - Residential dataset (55 homes)
- Prediction - Training
  - Validation-based early stopping
  - Building-based validation
  - Leave-*house*-out cross validation [2]
  - Complete (54 houses, 30 kHz, 11 cat.)
- Prediction - Evaluation
  - Category-based, F_1- score $F1S$
  - Building-based, Accuracy α
- Total accuracy: $\alpha = 0.89\%$



Three best/worst predicted categories

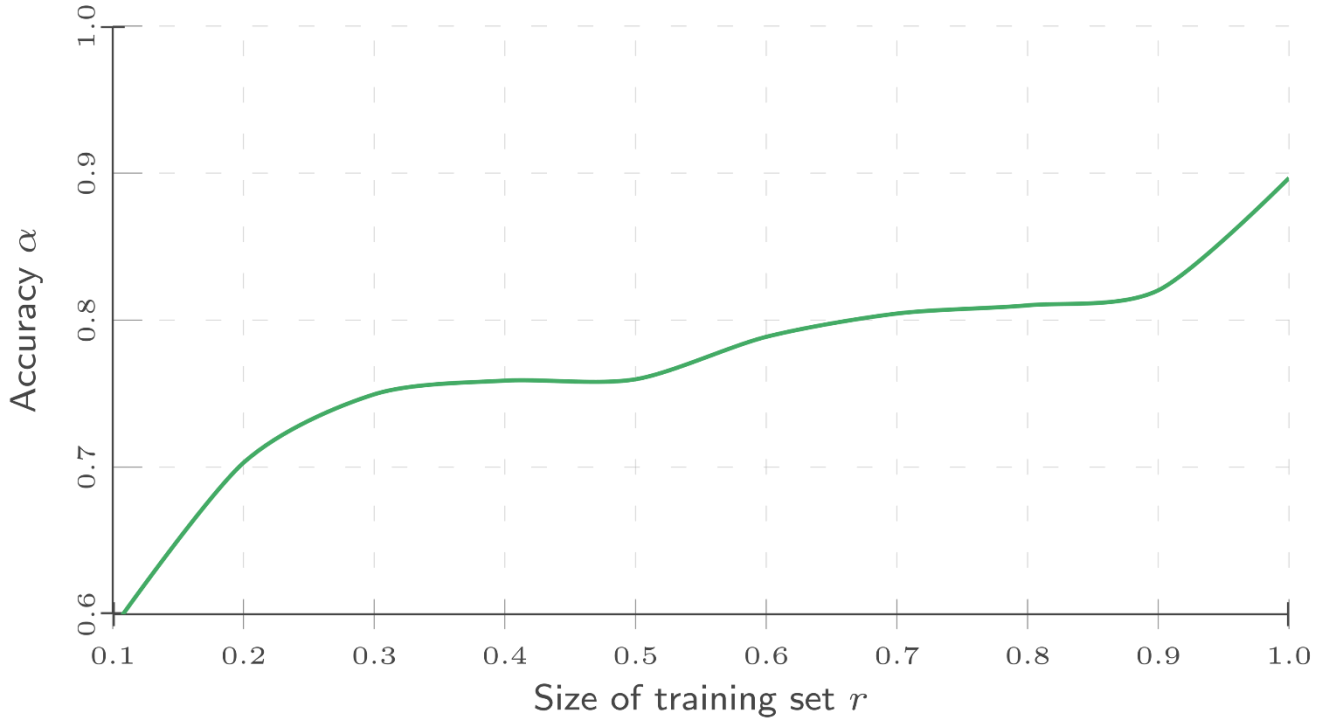
[2] J. Gao, E. C. Kara, S. Giri, and M. Berg´es, “A feasibility study of automated plug-load identification from high frequency measurements”. In proceedings of the 3rd IEEE Global Conference on Signal and Information Processing (GlobalSIP):220–224, Dec. 2015.

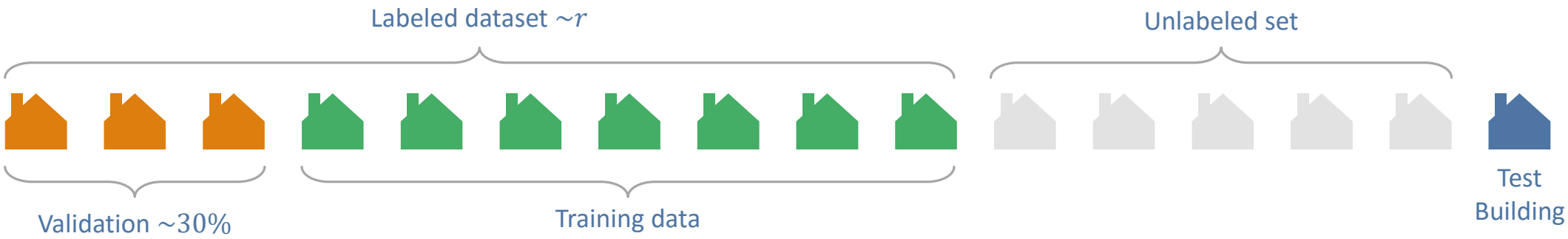
Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Experiments (2) - Training data

- **Effect of reducing size of training data r**
  - **Building-based reduction**
  - **Training on a ratio r of labeled data**
- Sampling frequency $f_s = 30$ kHz
- Test sample: last period of each measurement
- Label space $|\Omega| = 11$ categories

- Notable degradation w.r.t training data



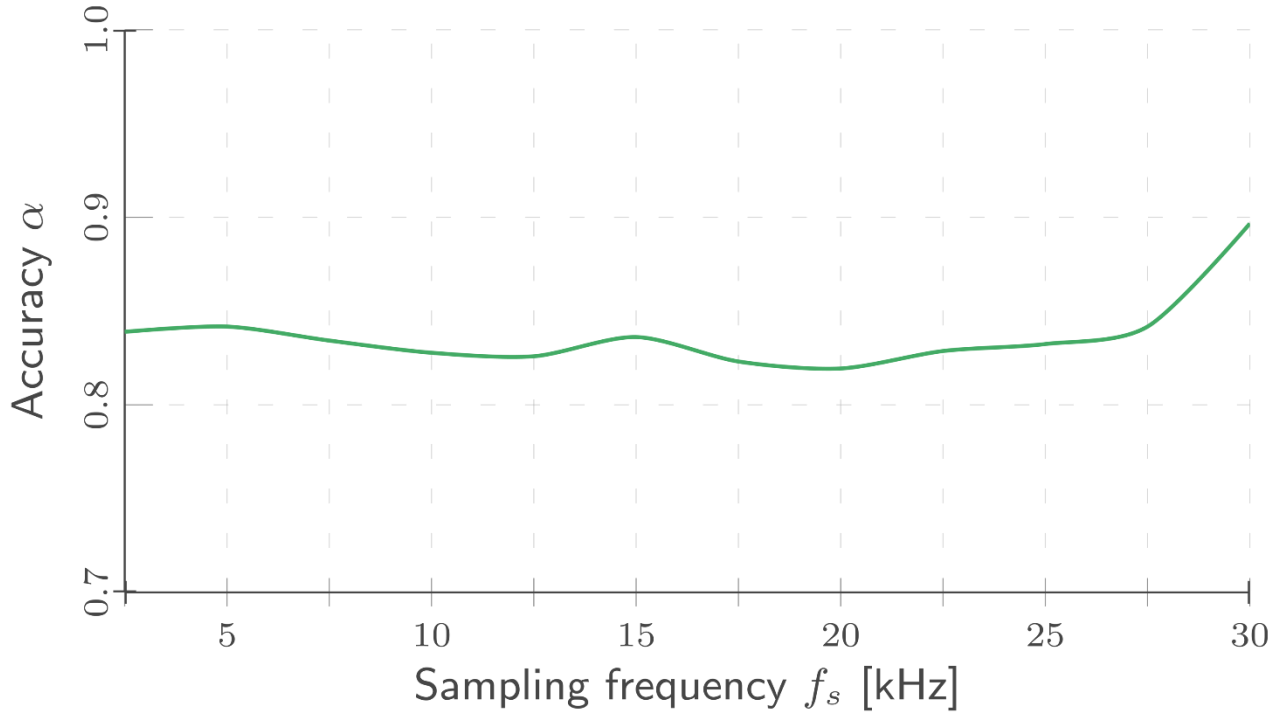


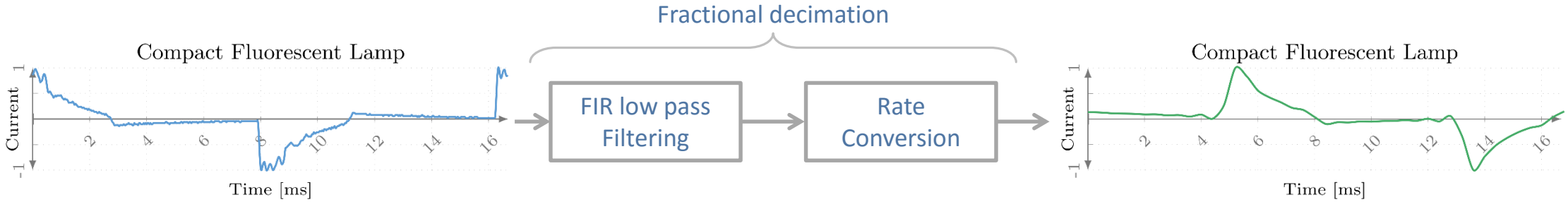
L. Mauch
Stuttgart University

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Experiments (3) - Sampling frequency

- Complete training data (54 buildings)
- **Resampled signals**
  - **Sampling frequency $f_s = 2.5 - 30$ kHz**
- Test sample: last period of each measurement
- Label space $|\Omega| = 11$ categories
- Almost stable for a wide range
  - Always +80%





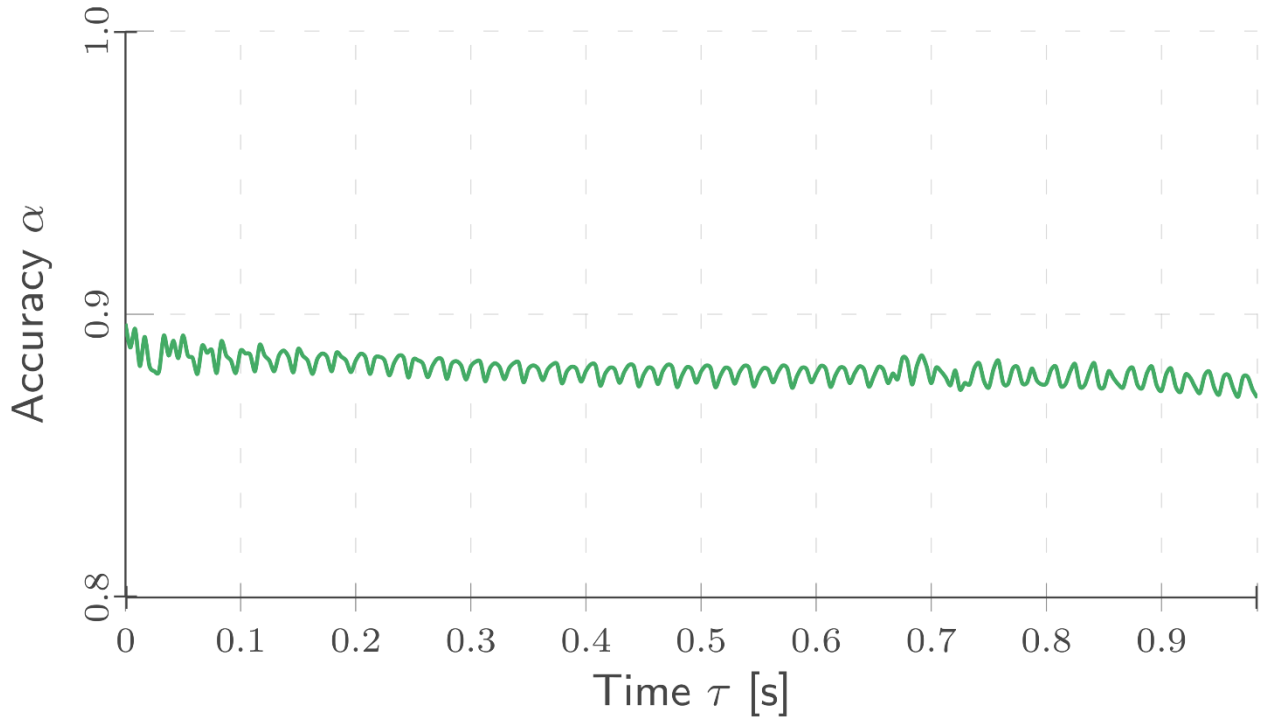
L. Mauch
Stuttgart University

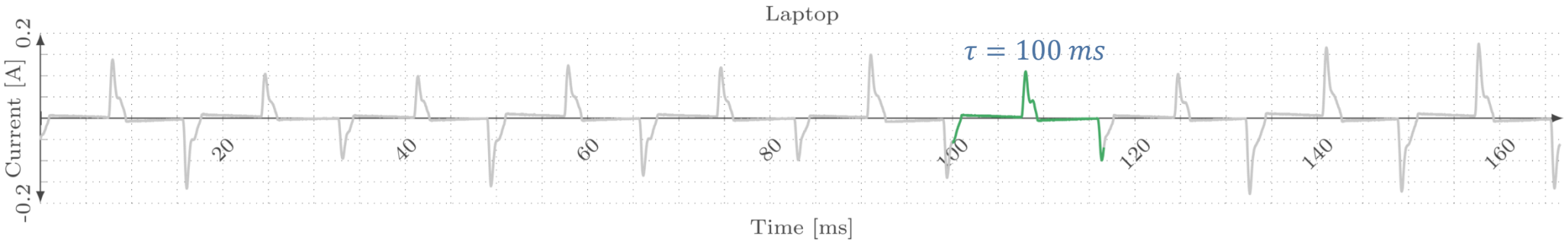
Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Experiments (4) - Signature variation

- Complete training data (54 buildings)
- Sampling frequency $f_s = 30$ kHz
- **Test sample:**
  - **Phase-shifted sample by τ seconds**
- Label space $|\Omega| = 11$ categories

- Robust to signature variations
  - But only evaluated on 1 second of operation
  - A certain phase shift is preferred



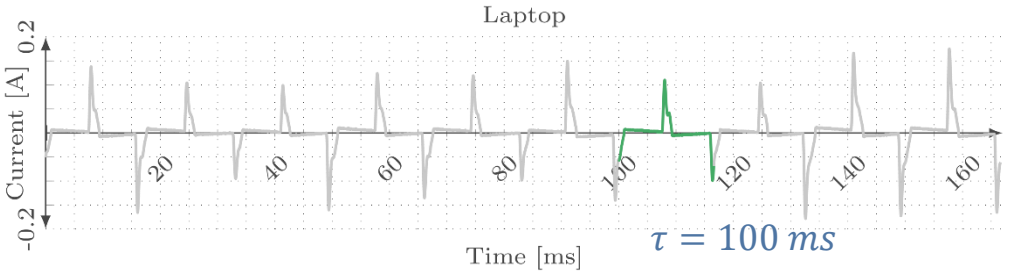


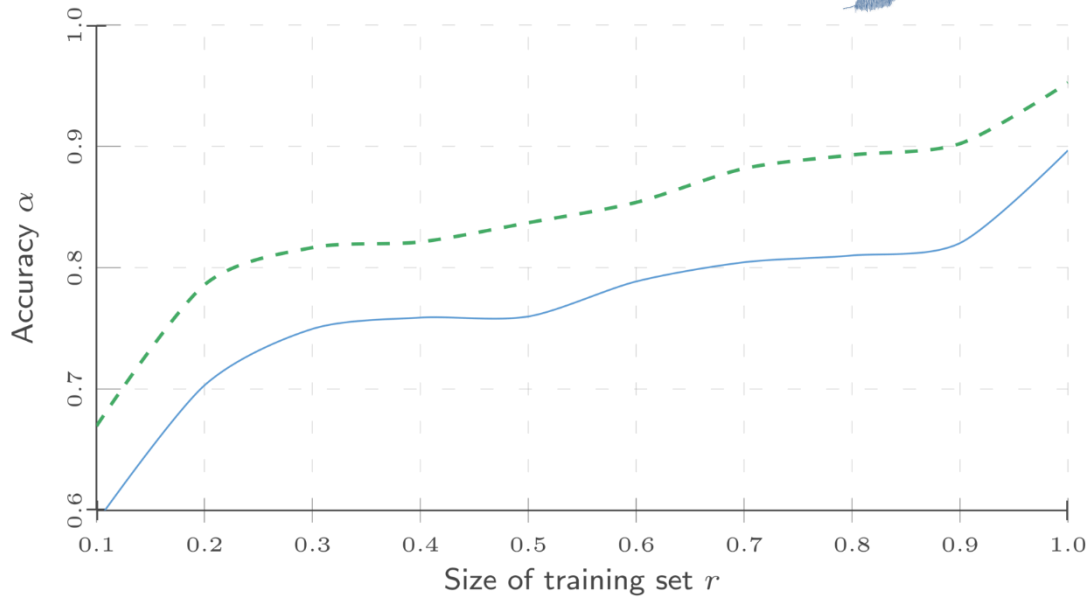
L. Mauch
Stuttgart University

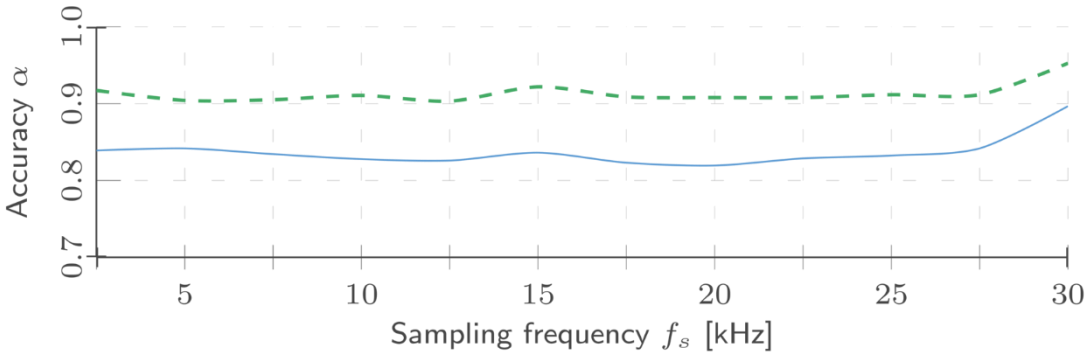
Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

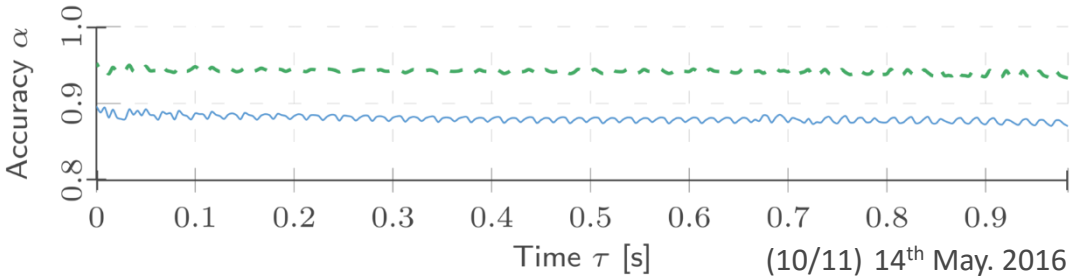
Experiments (5) - Label space

- Complete training data (54 buildings)
- Sampling frequency $f_s = 30$ kHz
- Test sample: last period of each measurement
- **Label space**
  - **A list of appliances in each building is known**
  - **A per-building label space** $\widehat{\Omega}_{\mathrm{h}}$
- Useful information but not always available.
  - Total accuracy: $\alpha \cong 95\%$









L. Mauch
Stuttgart University

Introduction
Task & tools
Signatures
Model
Prediction
Training data
Sampling frequency
Signature variations
Label space
Discussion

Discussion

Thank you
for your attention