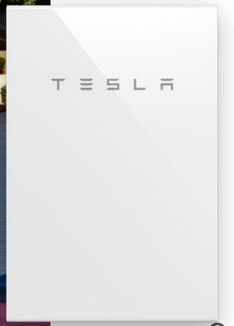
EnergyBoost: Learning-based Control of Home Batteries

Baihong Qi, Mohammad Rashedi, Omid Ardakanian
University of Alberta

ACM e-Energy 2019
June 2019
Phoenix, United States

This research has been undertaken thanks in part to funding from the Canada First Research Excellence Fund.

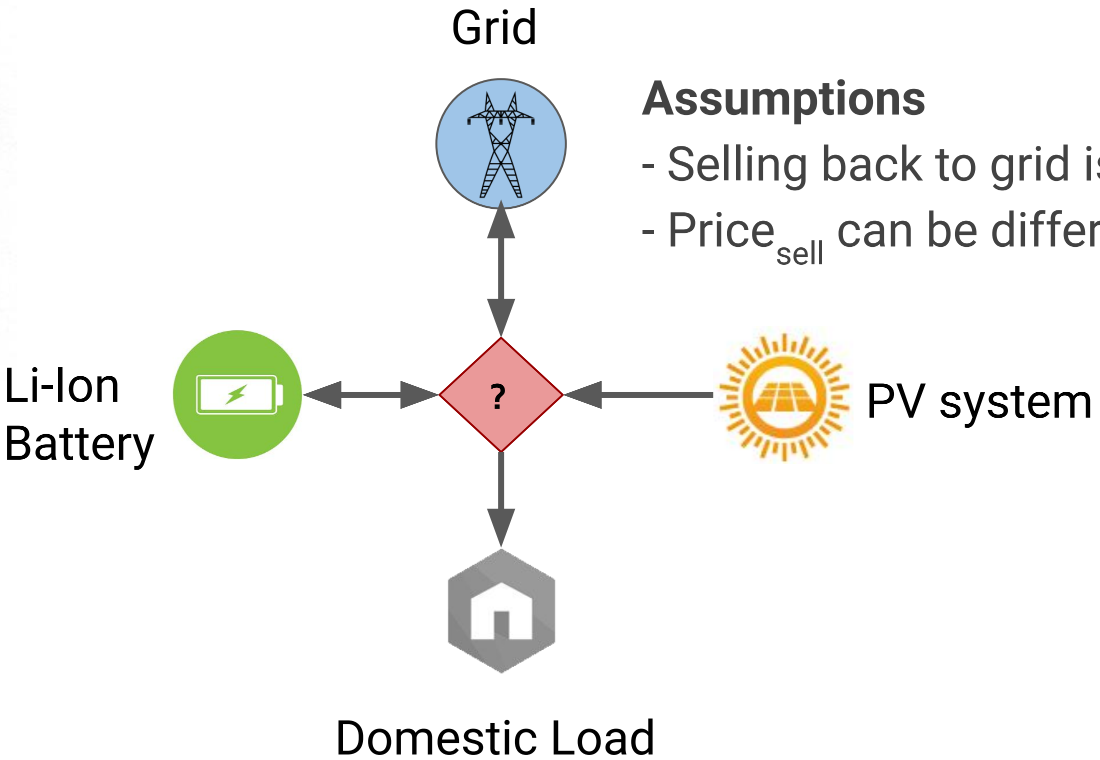
Solar is the fastest-growing source of renewable energy worldwide!



Tesla Powerwall

Photo by Gigaom: https://medium.com/@Scott_Wiener/lets-require-solar-panels-on-all-new-buildings-in-california-cb18fe9d9ec4

Controlling solar PV plus battery is nontrivial



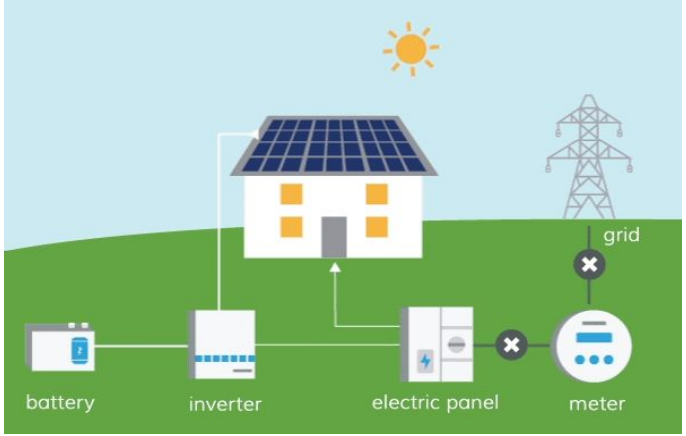
Assumptions

- Selling back to grid is permitted
- $\text{Price}_{\text{sell}}$ can be different from $\text{Price}_{\text{buy}}$

Basic idea...

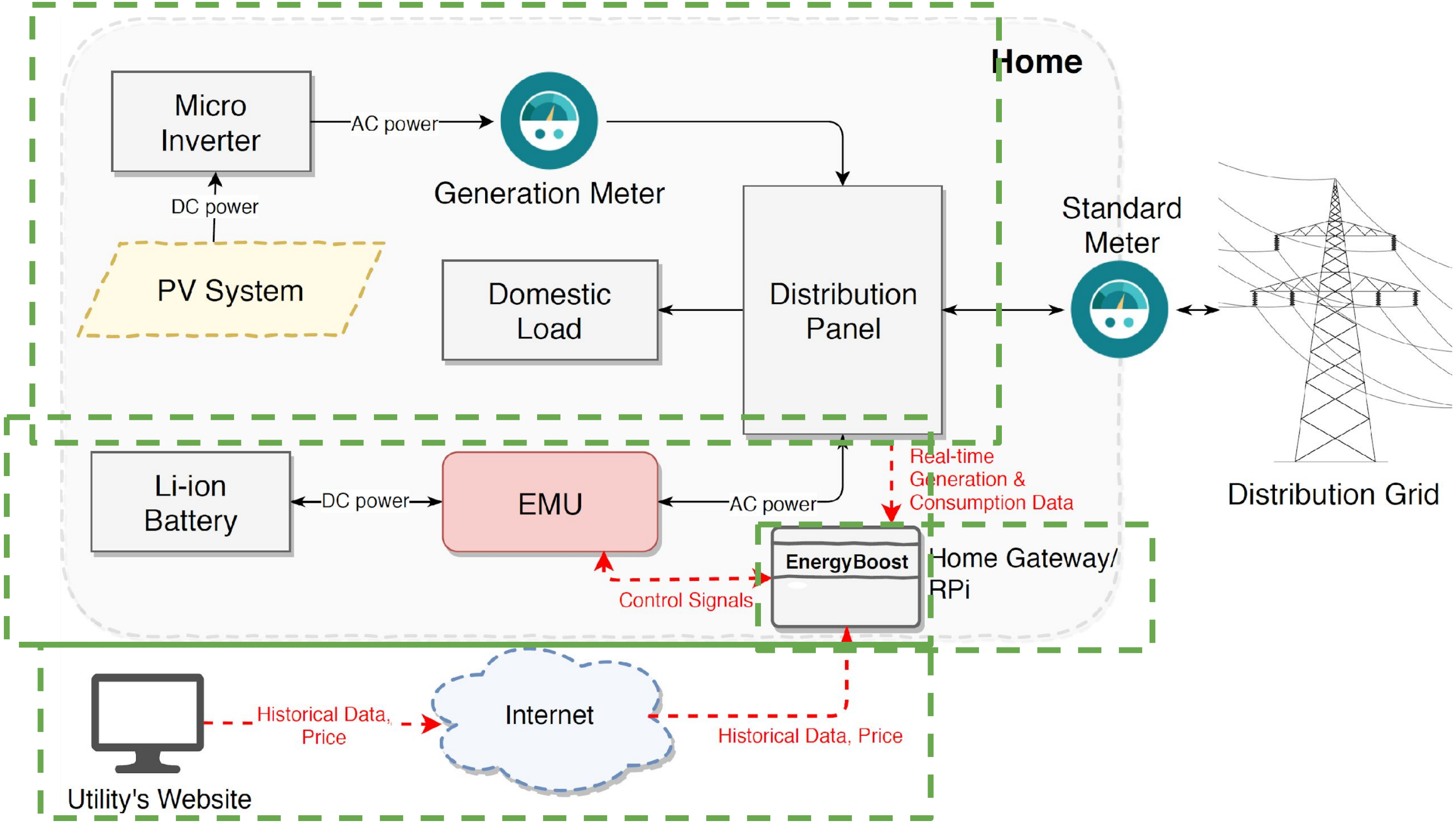
- EnergyBoost is a software program that runs on an edge device, automatically controlling **charge/discharge operations** of a battery

- It develops models based on historical data and leverages existing physics-based models to **minimize** the electricity bill of a customer over some time horizon



EnergyBoost

System architecture



Contributions

- ❑ The optimal control of battery is solved by defining and relaxing a mixed integer linear program (MILP)
- ❑ Using real household demands (70 homes) and solar generation traces the two predictive control policies implemented by EnergyBoost are compared with three baseline policies for different system sizes and solar tariffs
- ❑ The economic feasibility of battery-plus-PV is investigated through return on investment (ROI) analysis and break-even point calculations under a variety of scenarios

Mixed-integer linear program (MILP)

- The objective function is the net payment (**cost – revenue**) of a customer to the grid over some time interval

- Battery Operation constraints at each time interval are:
    - Battery has a finite capacity
    - Charge and discharge rates are bounded
    - Battery discharge rate must be less than the household demand
      i.e., cannot sell the stored energy to the grid
    - Buying energy from the grid and selling it back to the grid is not possible at the same time
    - Battery cannot be charged and discharged at the same time

- Integer variables are introduced to enforce the last two constraints

Predicting the future states

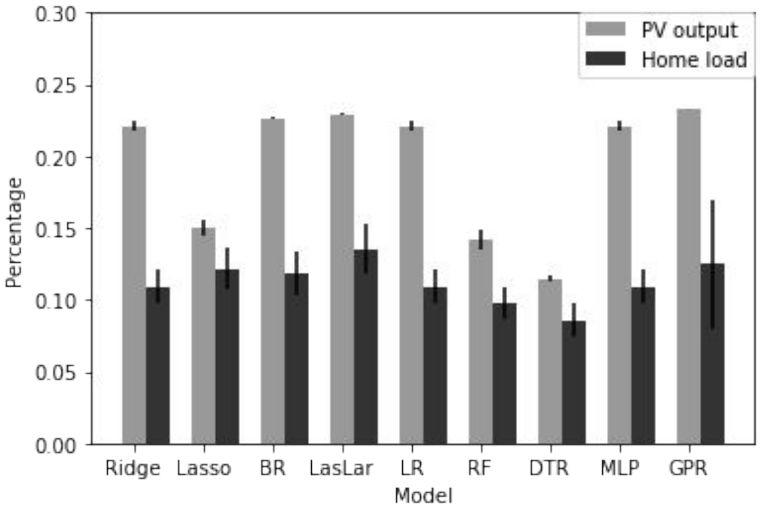
- Physics-based model
    - Linearized battery model from previous work that incorporates battery imperfections such as self-discharge and charge/discharge efficiency
    - Inverter’s model
- Data-driven model
    - Predictive model of the home demand
    - Predictive model of solar generation

Data-driven modelling

- Several features affect the home load and PV output
  Examples are weather data, historical home demand and PV output, and time

- ANOVA F-score is used to select the most relevant features

- The normalized RMSE of the next hour demand of home and PV output predictions for different models are compared



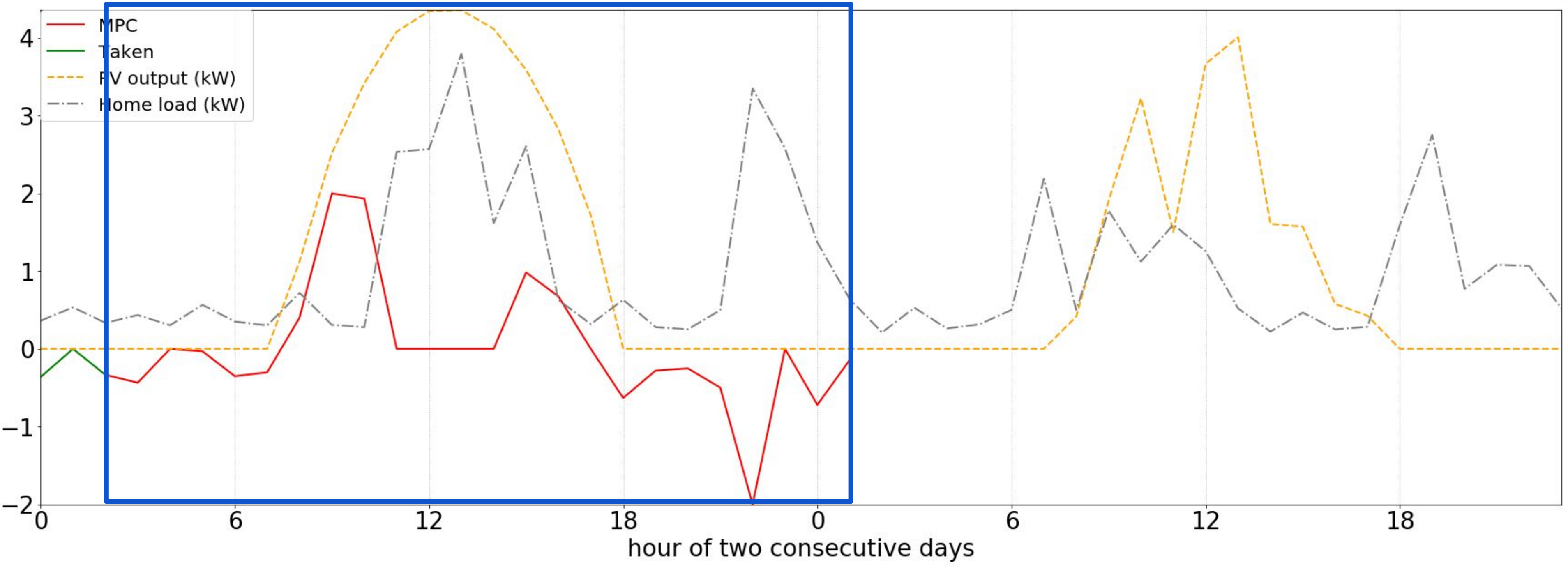
Error bars represent one standard error

Learning-based control methodology

- Model-based control
  - Model predictive control (MPC) solves a mixed integer linear program (MILP)

- Sample-based control:
  - Model-free control using Reinforcement Learning: Advantage Actor Critic Method

Model predictive control



Reinforcement learning

- An RL agent learns successful strategies from interactions (in a number of episodes) with the environment and discovers the action that maximizes the expected cumulative reward given the current state

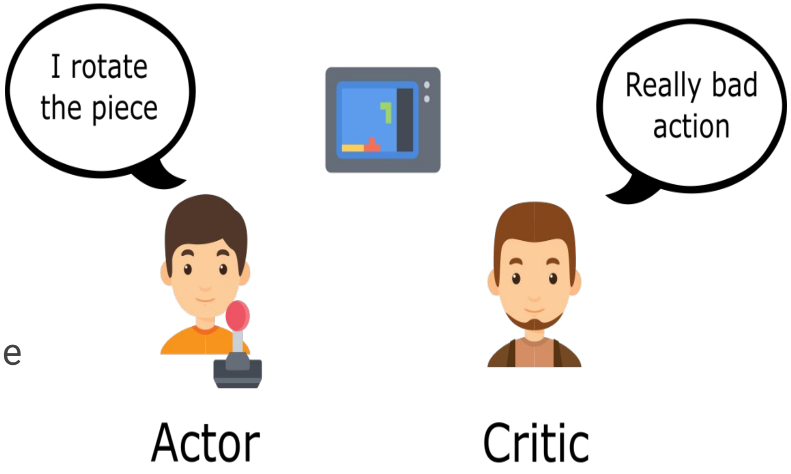
- Since these interactions can damage the battery we **simulate** these interactions

- The immediate reward is defined as the payment for the current time slot

- **Challenges**: Continuous control (continuous state/action space)
  Time-varying constraints

Actor-critic method

- **Actor** controls how our agent behaves (policy-based)
- **Critic** measures how good is the action taken by the actor (value-based)
- The actor-critic method makes an update in every step using Q(s,a)



Policy Update: $\Delta\theta = \alpha \, \nabla_\theta(log \; \pi_\theta(s,a))\underbrace{\hat{q}_w(s,a)}_{}$

Policy Evaluation

q learning function approximation (estimate action value)

Value update: $\Delta w = \beta \; \underbrace{(R(s,a) + \gamma \hat{q}_w(s_{t+1}, a_{t+1}) - \hat{q}_w(s_t, a_t))}_{\text{TD error}} \underbrace{\nabla_w \hat{q}_w(s_t, a_t)}_{\text{Gradient of our value function}}$

Policy and value have different learning rates

Data Sets

- **Data port**: Household related data
  - Home use, temperature, cloud cover, wind speed
  - 70 households in Austin, Texas
  - 15-minute data
- **Solar Research Lab**: Global horizontal Irradiance in Texas
  - Hourly data
- **ComEd**: Hourly Prices in Illinois
- **Power stream**: Time of Use prices in Ontario
  - Hourly data

Simulation scenarios

To evaluate the control algorithms more effectively, several scenarios are considered:

- ❏ Three battery sizes: 0, 6.4kWh/2kW, and 13.5kWh/5kW

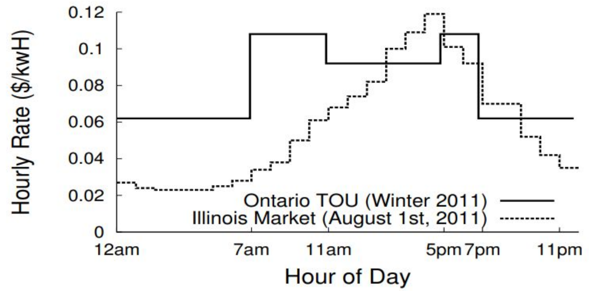
- ❏ Three solar panel sizes: 0, 4.4kWp, and 8.8kWp

- ❏ Two pricing schemes: TOU and hourly

- ❏ Four solar export tariffs: 3, 6.1, 7.7, and 15.4 cents/kWh

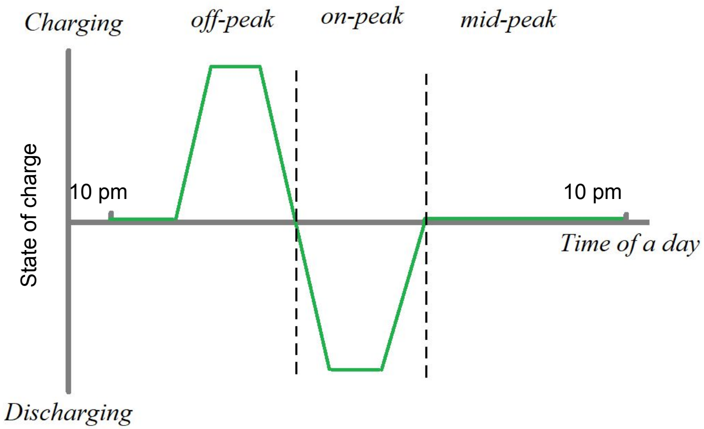
TOU price	7am-11am	11am-5pm	5pm-7pm	7pm-7am
Nov. to Apr.	0.101	0.072	0.101	0.05
May to Oct.	0.072	0.101	0.072	0.05

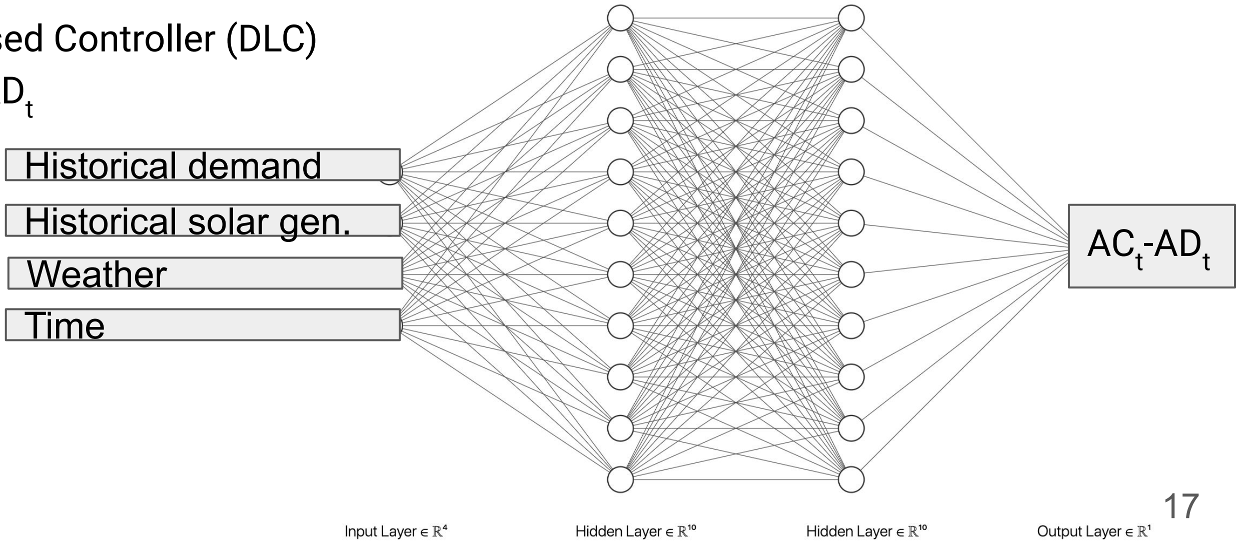
- *Off-peak*: when the cost and demand are low
- *Mid-peak*: when the cost and demand are moderate
- *On-peak*: when the cost and demand are high



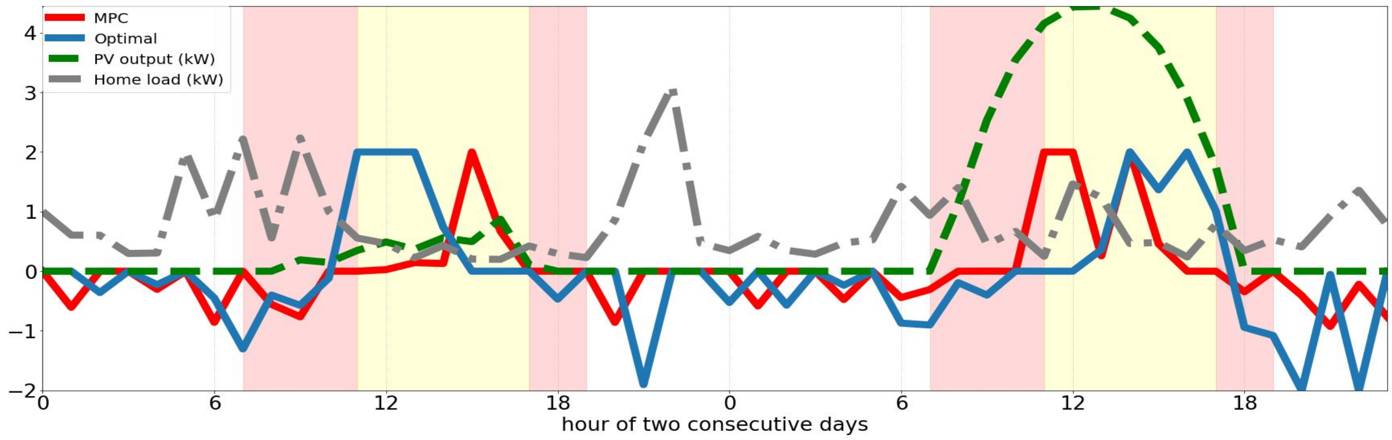
Baseline policies

- ❏ MILP with Oracle (Optimal)
- ❏ Rule-based controller (RBC)
- ❏ Direct Learning-based Controller (DLC)
  $F(\textit{features}_t) = AC_t\text{-}AD_t$





How does the learned strategy look like?



A sample home with

- 4.4kWp PV system
- Tesla Powerwall 1
- solar export tariff of 0.03$/kWh.

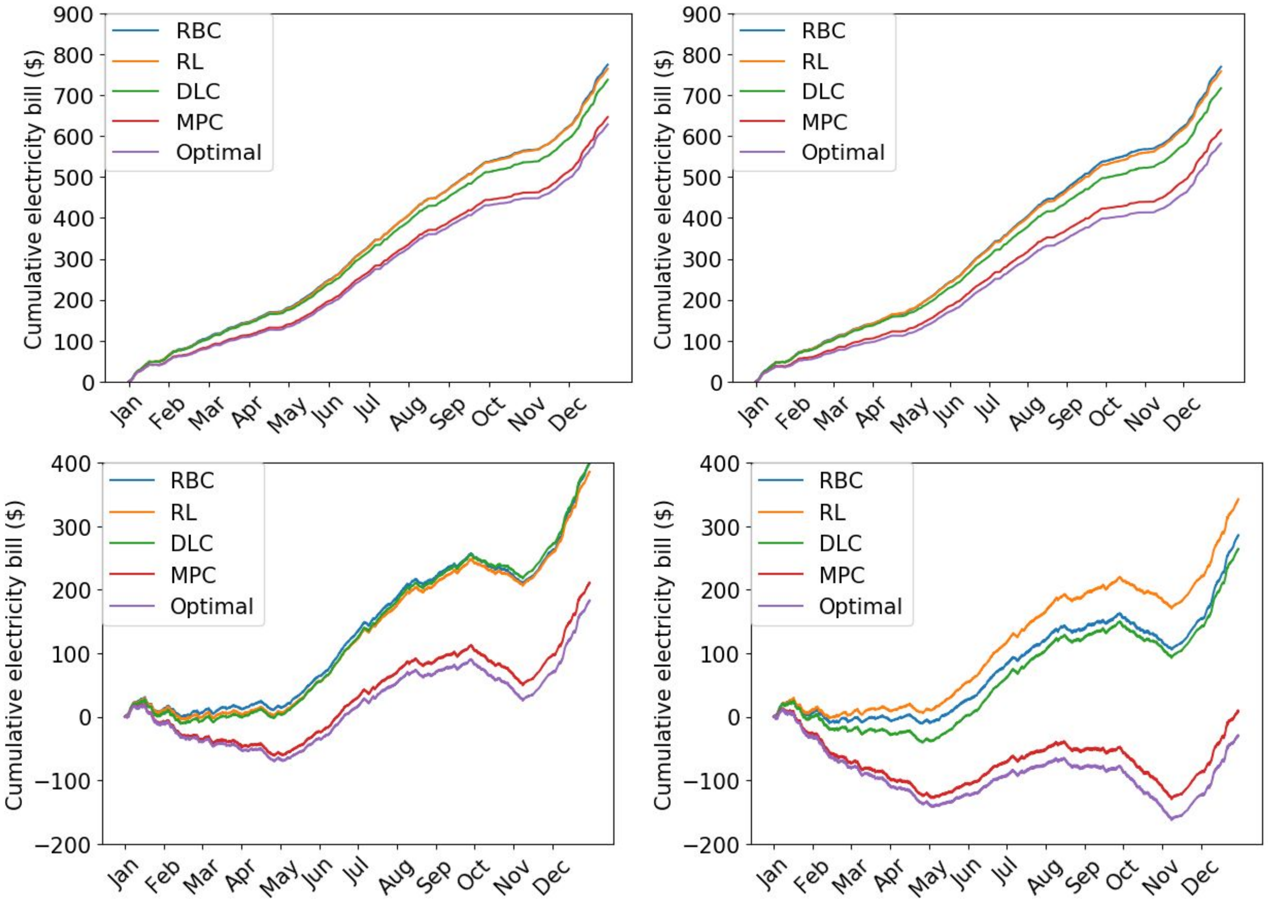
The on-peak and mid-peak intervals are highlighted in red and yellow

Annual Bills

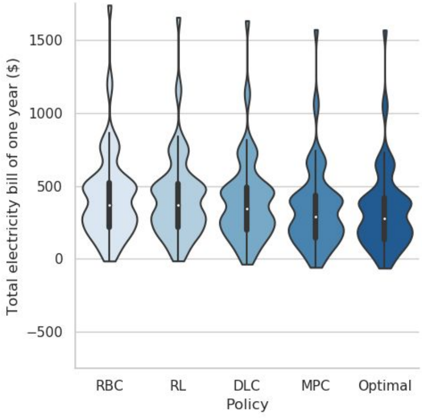
A randomly selected home with a 4.4kWp PV system

Tesla Powerwall battery (left column: 6.4 kWh; right column: 13.5 kWh).

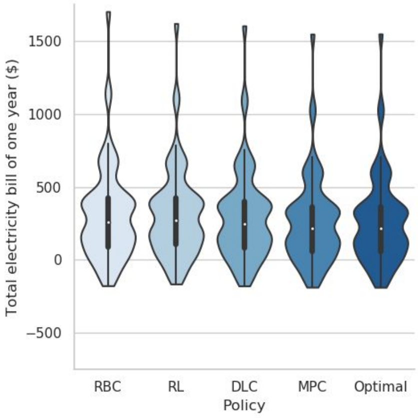
The solar tariff is
 0.03$/kWh (top row) and
0.154$/kWh (bottom row)



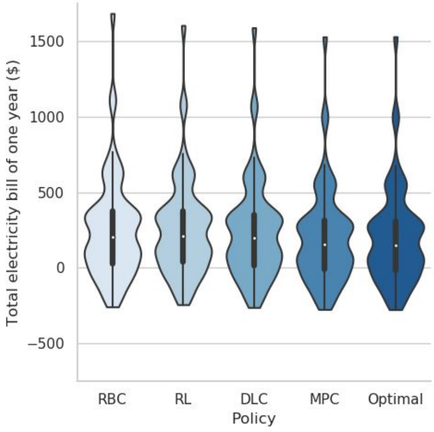
Which control policy performs better?



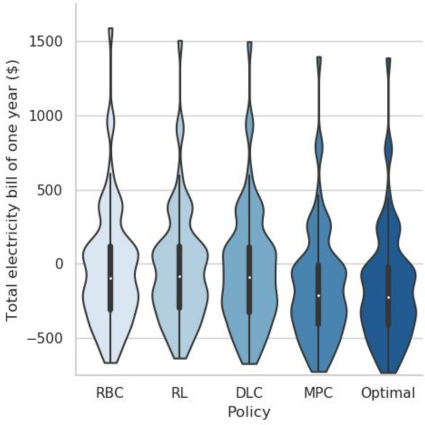
Solar tariff = 0.03$/kWh



Solar tariff = 0.061$/kWh



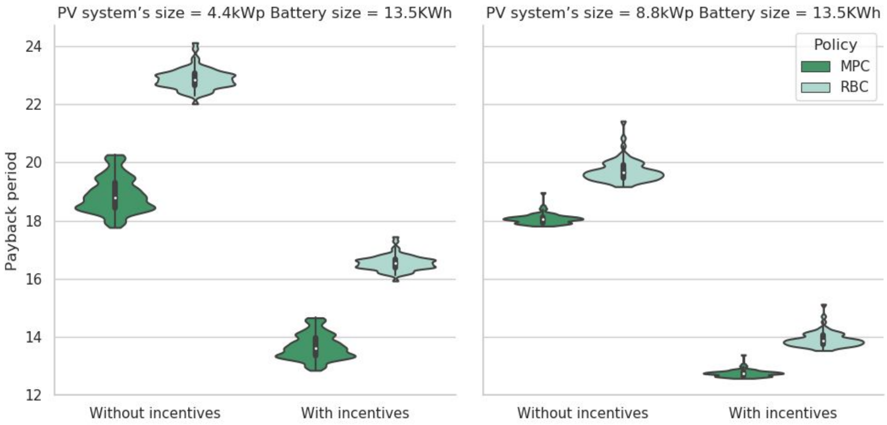
Solar tariff = 0.077$/kWh



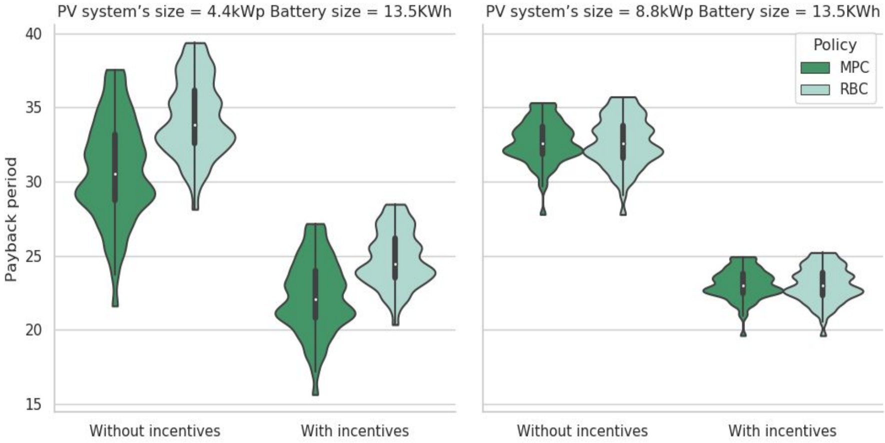
Solar tariff = 0.154$/kWh

Annual electricity bill of homes equipped with a 4.4kWp PV system and a Tesla Powerwall~1 under TOU pricing scheme.

Break even point



Solar tariff = 0.154$/kWh



Solar tariff = 0.077$/kWh

certain system sizes are profitable in 20 years with existing incentives under this tariff structure

Conclusions

- EnergyBoost utilizes learning-based control strategies to determine optimal battery operations

- The best learning-based controller (i.e., MPC) outperforms baseline controllers in terms of the annual electricity bill

- EnergyBoost reduces the payback period by more than 22 months on average compared to the rule-based controller

- Only some battery sizes are profitable (in 20 years) under some tariff structures