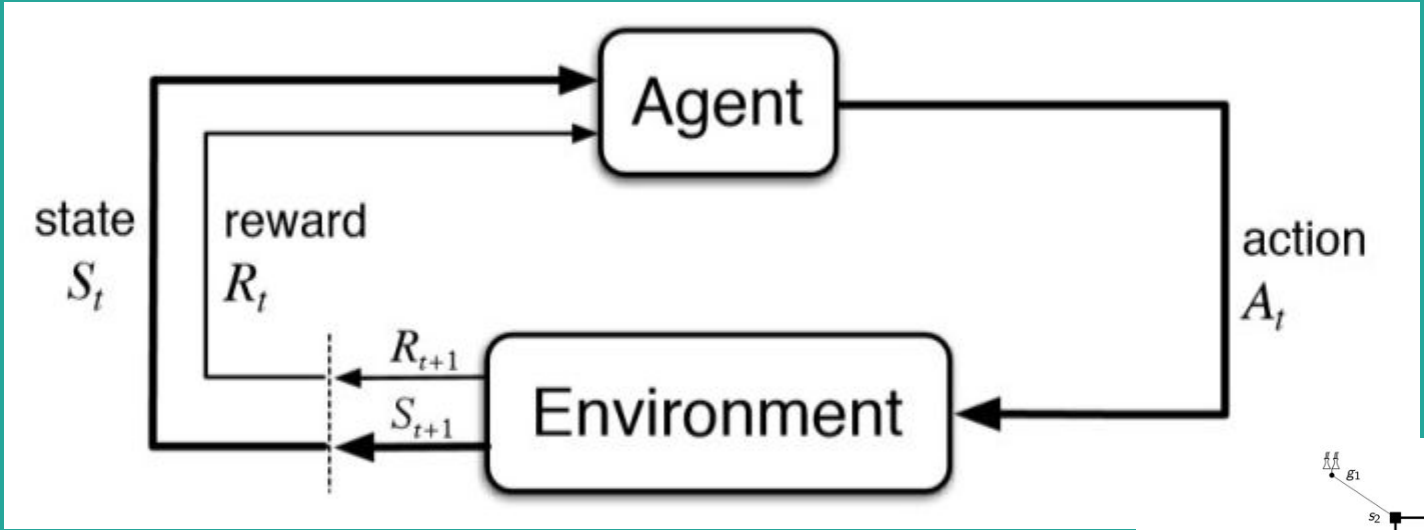
Learning To Run A Power Network

Turing AI Fellows and Teams Community Hackathon

Avirup Das, Farajollah Tahernezhadjavazm, Filippo Valdettaro,
Satya Prakash Dash, Taku Yamagata

Reinforcement Learning



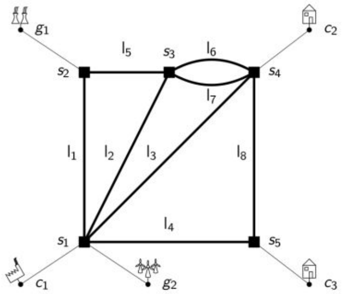


Figure 3: A simple electricity network, showing the circuit nature of a power network, the currents I flowing in the lines and the *interconnectedness* between generators denoted g, customer loads denoted c and substation nodes denoted s.

Constraints

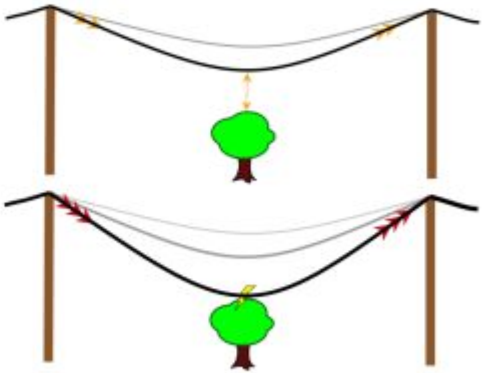


Figure 2: An example of the dangers of overheating power-lines, by transporting too much current, the metallic conductor heats and sags close to the ground causing a flash over to ground and endangering human life.

ρ_i: current flow / thermal limit

$\rho_i < 1.0$

Reinforcement Learning with Constraints

IPO (Interior Point Optimization)[1]

IPO augments the objective function with logarithmic barrier functions as penalty functions to accommodate the constraints.

$$\max_{\theta} J_R^{\pi_\theta} + \sum_{i}^{n\text{-}line} \frac{1}{\mathrm{t_i}} \log(1.0 - \rho_i)$$

Accumulated Reward
(objective function)

$\mathrm{t_i}$: hyper-parameter (scaling the regularizer term)

ρ_i: current flow / thermal limit for i-th line

[1] Yongshuai Liu, Jiaxin Ding, and Xin Liu. Ipo: Interior-point policy optimization under constraints. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, pages 4940–4947, 2020

Observations

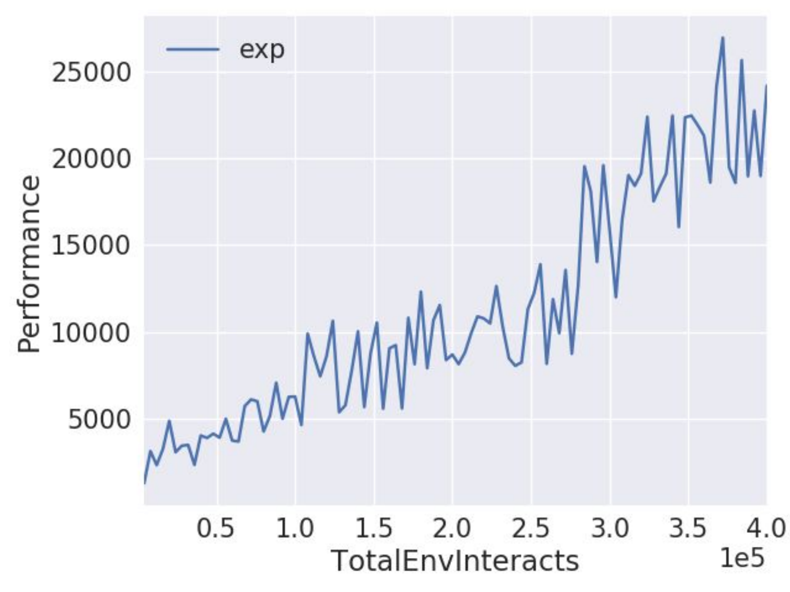
- All continuous observations have been taken into account.
- Observations (active and reactive power, voltage and other parameters) from generator and loads.
- Topology of the grid.
- Parameters from storage.

Actions

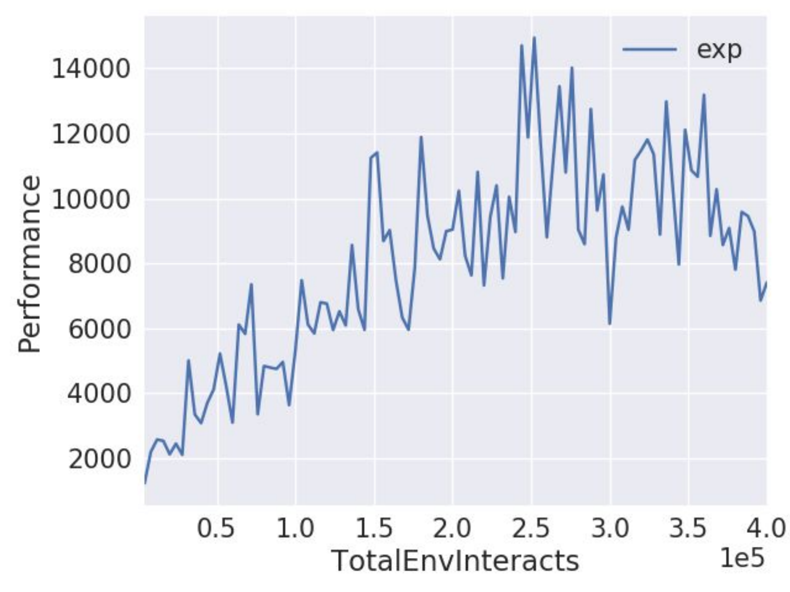
- All discrete actions.
- Change or Set busbars for generators, loads, lines and storage.

Training till Saturation

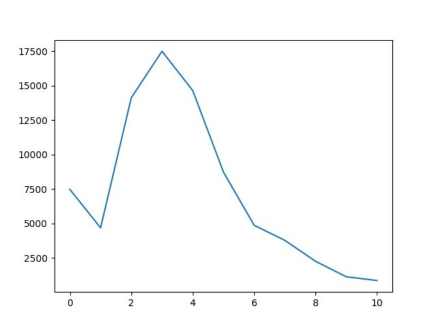
With regularisation

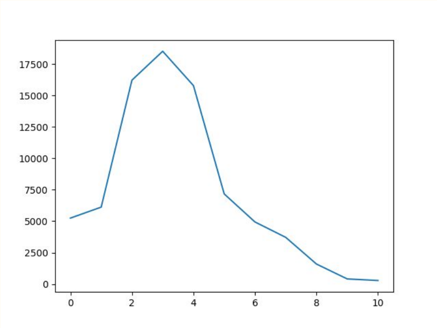


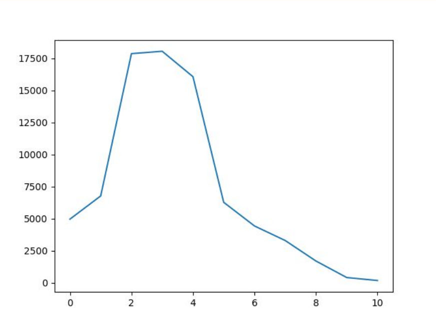
Without regularisation

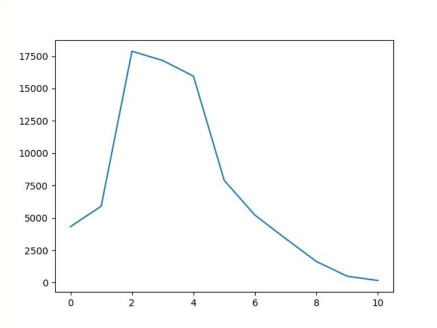


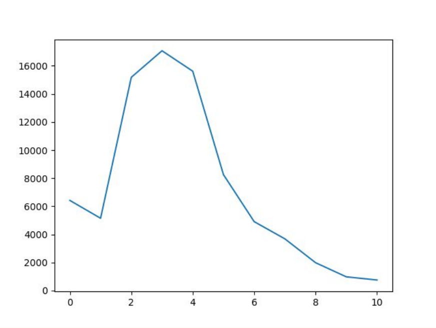
Distribution of ρ over time

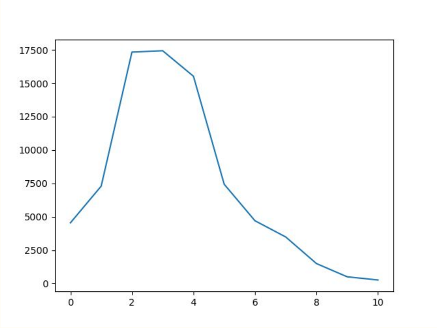


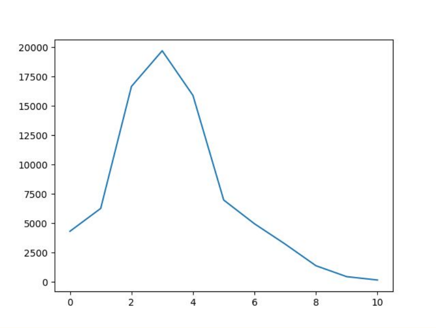


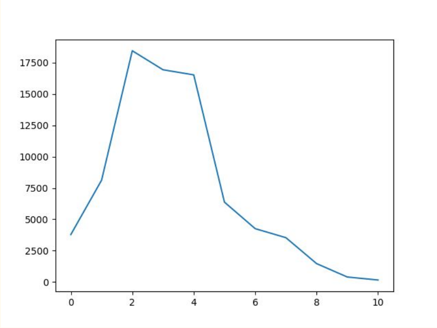






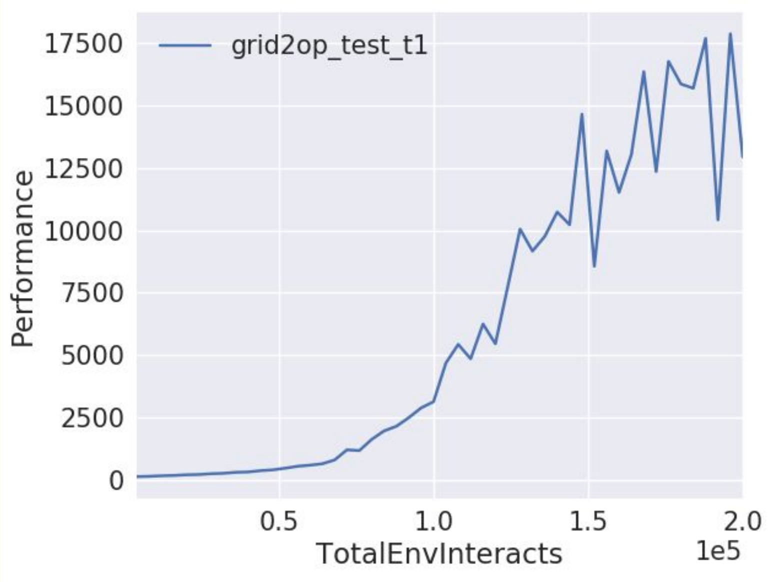




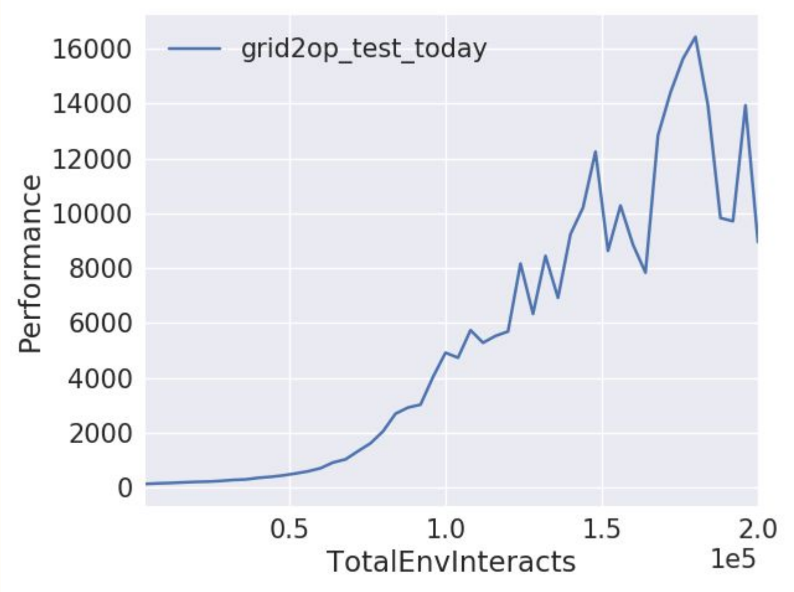


Learning curve

With constraint penalty (t=1.0)



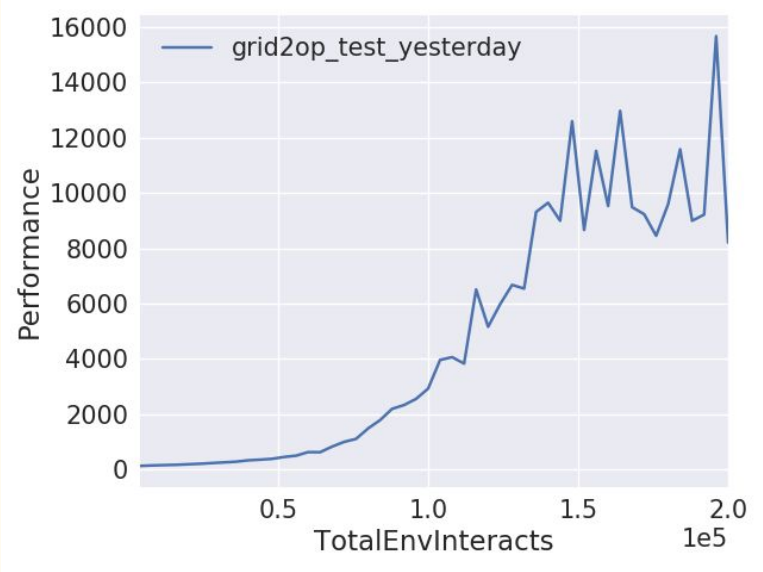
No constraint penalty



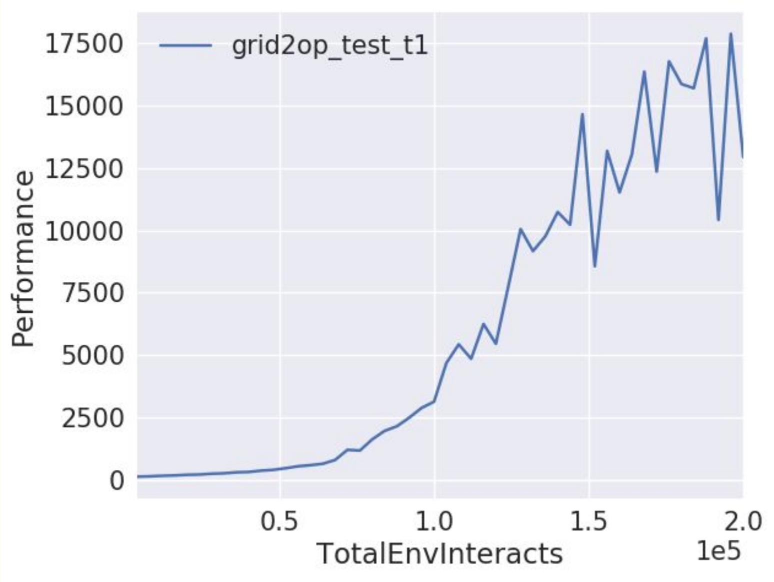
Do nothing baseline: approximately around 3000

Hyperparameter (t)

t=0.1

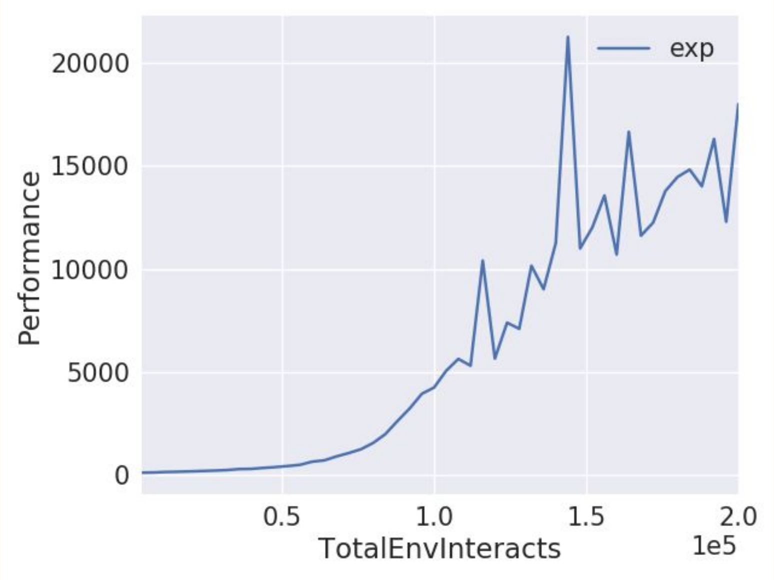


t=1.0

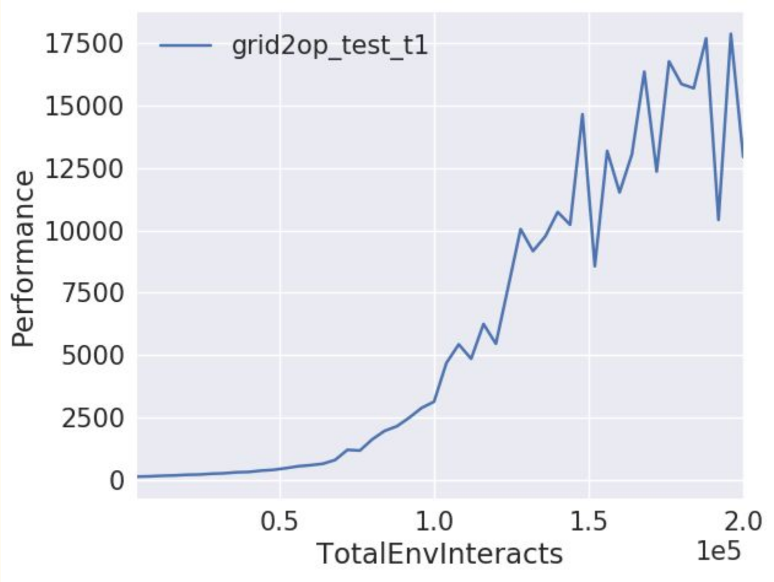


Hyperparameter (t)

t=0.5



t=1.0



Conclusion

Penalty constraints can help