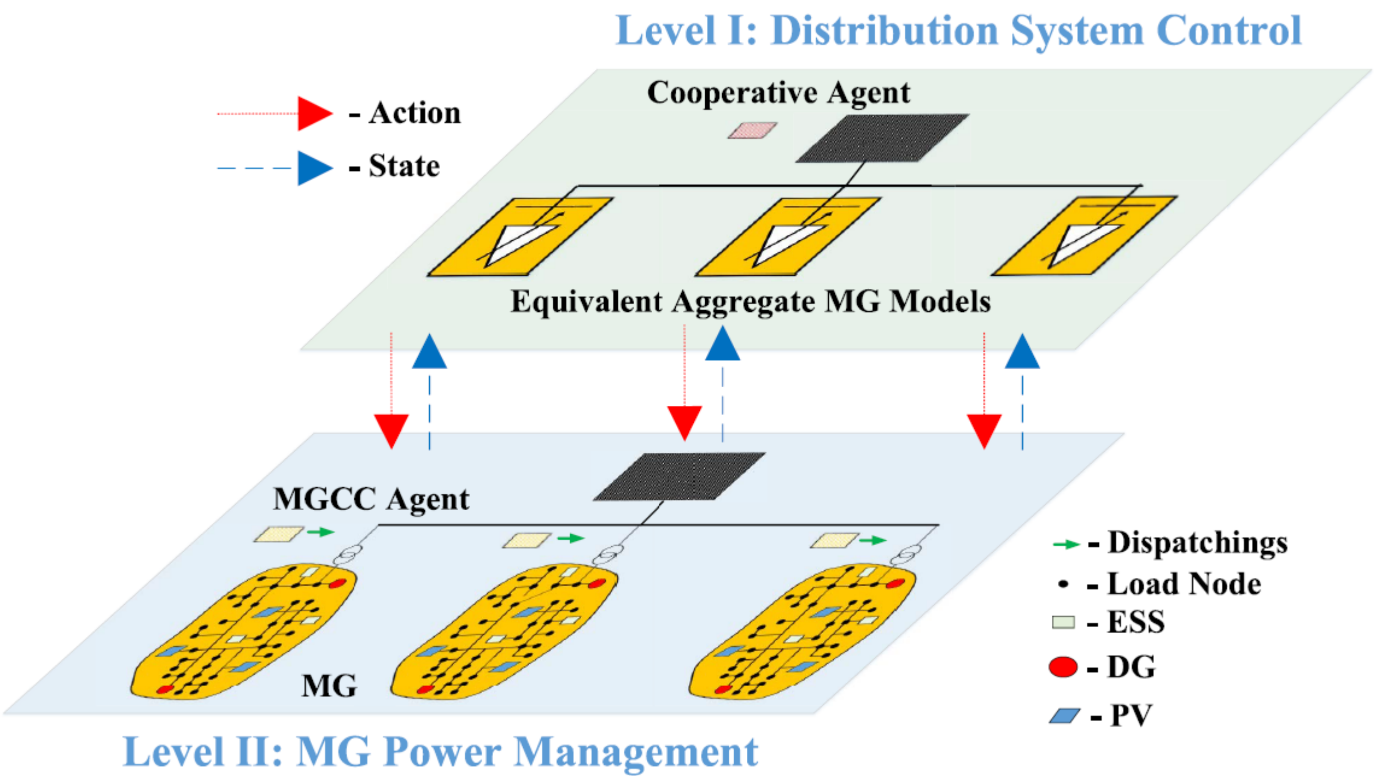
RL-based Energy Management of MGs



A bi-level RL-based energy management of networked MGs under incomplete information

- At the Level I, **a cooperative RL agent** performs function approximation to predict the behavior of MGs and give retail price signals for MGs.
- At the Level II, the **MG agents** receive the price signals from the Level I then, each MG provides power-flow-constrained optimal responses to price singles.
- Only uses **aggregate data** at the point of common coupling

Q. Zhang, K. Dehghanpour, Z. Wang and Q. Huang, "A Learning-Based Power Management Method for Networked Microgrids Under Incomplete Information," in IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1193-1204, March 2020.

RL-based Energy Management of MGs

Level I: one cooperative agent solves a RL problem:

- Maximize the reward function

$$R(t) = \sum_{t'=1}^{T-1} \gamma^{t'} \left(\lambda^{W}_{t+t'} P^{W}_{t+t'} - \sum_{n=1}^{N} \lambda^{R}_{t+t',n} P^{PCC}_{t+t',n} \right)$$

- The state-action value function (Q-function)

$$Q^{*}_{t}(\boldsymbol{S},\boldsymbol{a}) = E\left\{\pi(t+1) + \gamma \cdot \max_{\boldsymbol{a}'} Q^{*}_{t}(\boldsymbol{S}(t+1),\boldsymbol{a}')\right\}$$

- Parameterize the Q-function with multivariate regression method

$$Q_t(\boldsymbol{S},\boldsymbol{a}) \approx \hat{Q}_t(\boldsymbol{S},\boldsymbol{a}|\boldsymbol{\theta}) = Q_{\boldsymbol{S}\cdot\boldsymbol{a}}(t|\boldsymbol{\theta}) + Q_{\boldsymbol{S}}(t|\boldsymbol{\theta}) + Q_{\boldsymbol{a}}(t|\boldsymbol{\theta})$$

- Greedy action selection

$$\boldsymbol{a}_{opt}(t') = \ arg \max_{\boldsymbol{a}'} Q_{t'}(\boldsymbol{S}(t'),\boldsymbol{a}')$$

- Update regression parameter

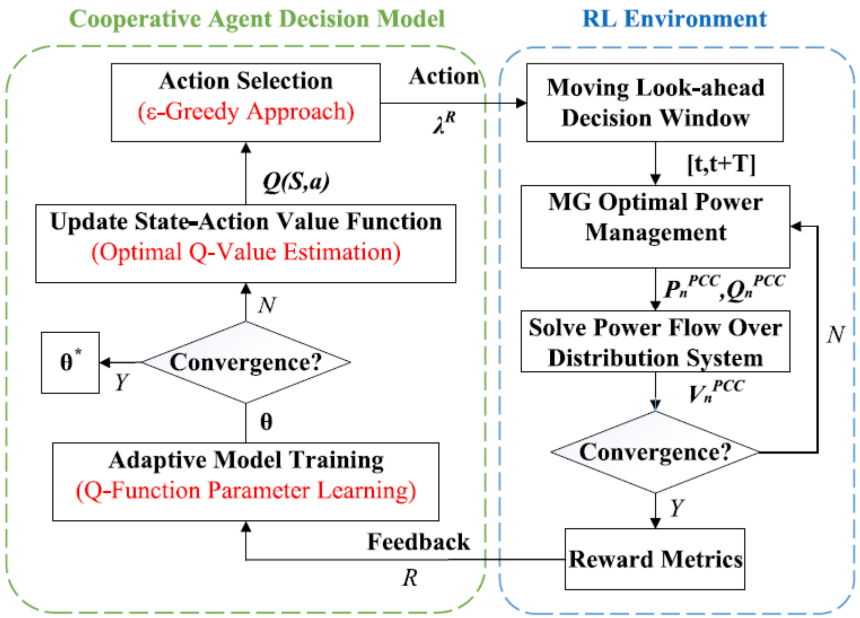
$$\theta(t+1) \leftarrow \theta(t) + \Delta(t)x(t)\{R(t) - \hat{Q}_t(\boldsymbol{S},\boldsymbol{a}|\boldsymbol{\theta})\}$$

$$\Delta(t+1) \leftarrow \hat{\Delta}(t+1)\left(I + \mu\hat{\Delta}(t+1)\right)^{-1}$$

$$\hat{\Delta}(t+1) \leftarrow \frac{1}{1-\phi}\left(\Delta(t) - \frac{\Delta(t)x(t)x^{T}(t)\Delta(t)}{1 + x^{T}(t)\Delta(t)x(t)}\right)$$

Action of utility agent:

- Locational retail price



Reward and State of MGs:

- Aggregated information of load and DERs

Level II: each MG solves a constrained optimal power flow problem problem:

- Minimize operational cost of each MG

$$\min_{x_p,x_q} \sum_{t}^{T+t} \left(-\lambda^{R}_{t,n} P^{PCC}_{t,n} + \lambda^{F}_{i,t,n} F_{i,t,n}\right)$$

- Power transfer between MGs

$$\left|P^{PCC}_{t,n}\right| \leq P^{PCC,M}_{t,n}$$

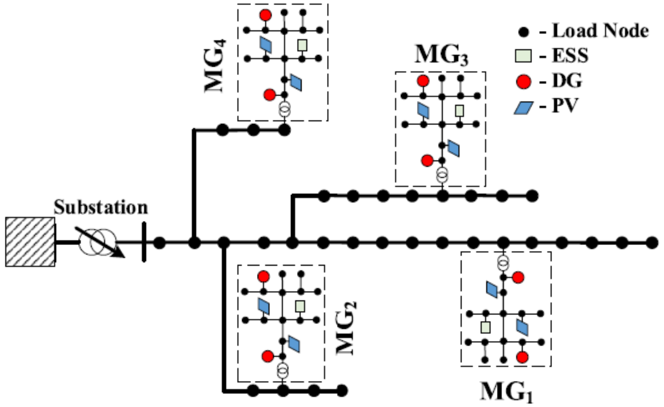
$$\left|Q^{PCC}_{t,n}\right| \leq Q^{PCC,M}_{t,n}$$

- Network OPF-based constraints
- Operational constraints of DGs
- Operational constraints of ESSs

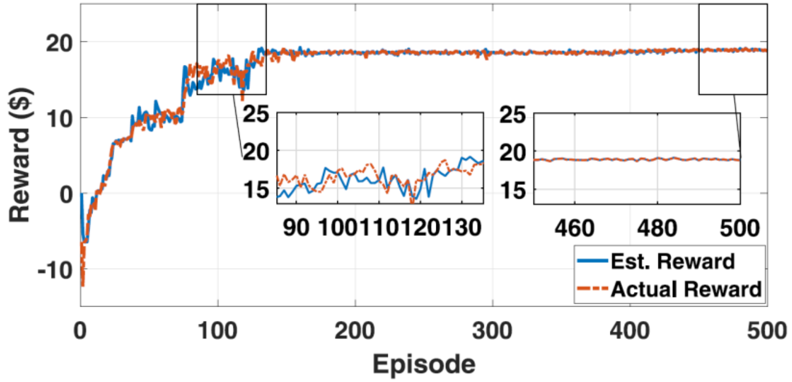
Q. Zhang, K. Dehghanpour, Z. Wang and Q. Huang, "A Learning-Based Power Management Method for Networked Microgrids Under Incomplete Information," in IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1193-1204, March 2020.

RL-based Energy Management of MGs

- Test distribution system with networked MGs



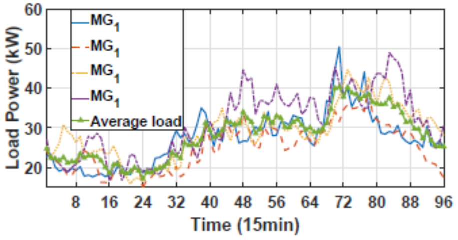
- Training performance



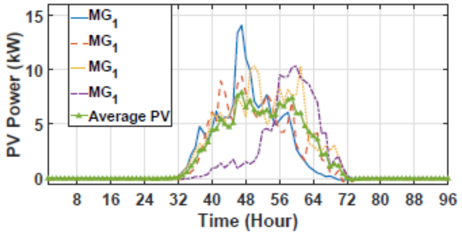
- Comparison between benchmark model-based methods and model-free methods

	RL-based method	Centralized Opt
Social welfare ($)	4212.372	4232.264
Computational time (s)	9.64	116.35
MG privacy maintenance	Yes	No

- Training data (4 year and 15-min smart meter data)

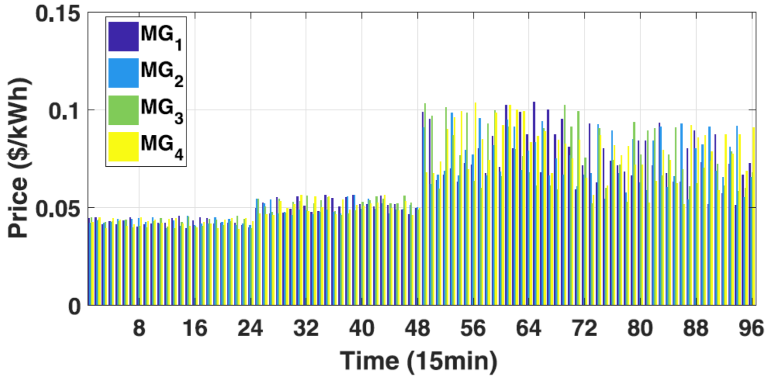


(a) Aggregate active load profile of the MGs

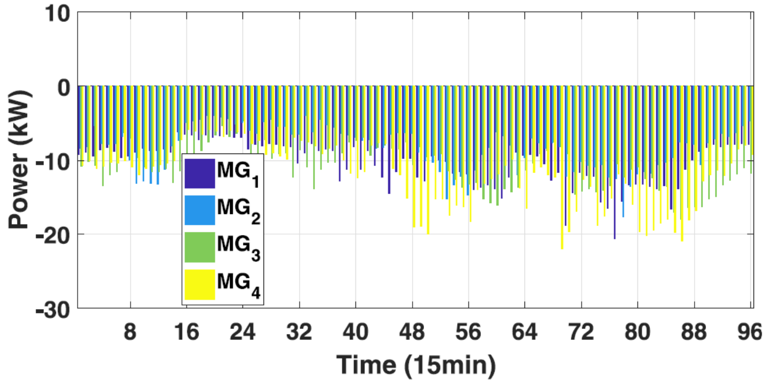


(b) Aggregate PV power of the MGs

- Retail price signals (Level I actions)



- Power transfer between MGs (Level II responses)



Q. Zhang, K. Dehghanpour, Z. Wang and Q. Huang, "A Learning-Based Power Management Method for Networked Microgrids Under Incomplete Information," in IEEE Transactions on Smart Grid, vol. 11, no. 2, pp. 1193-1204, March 2020.