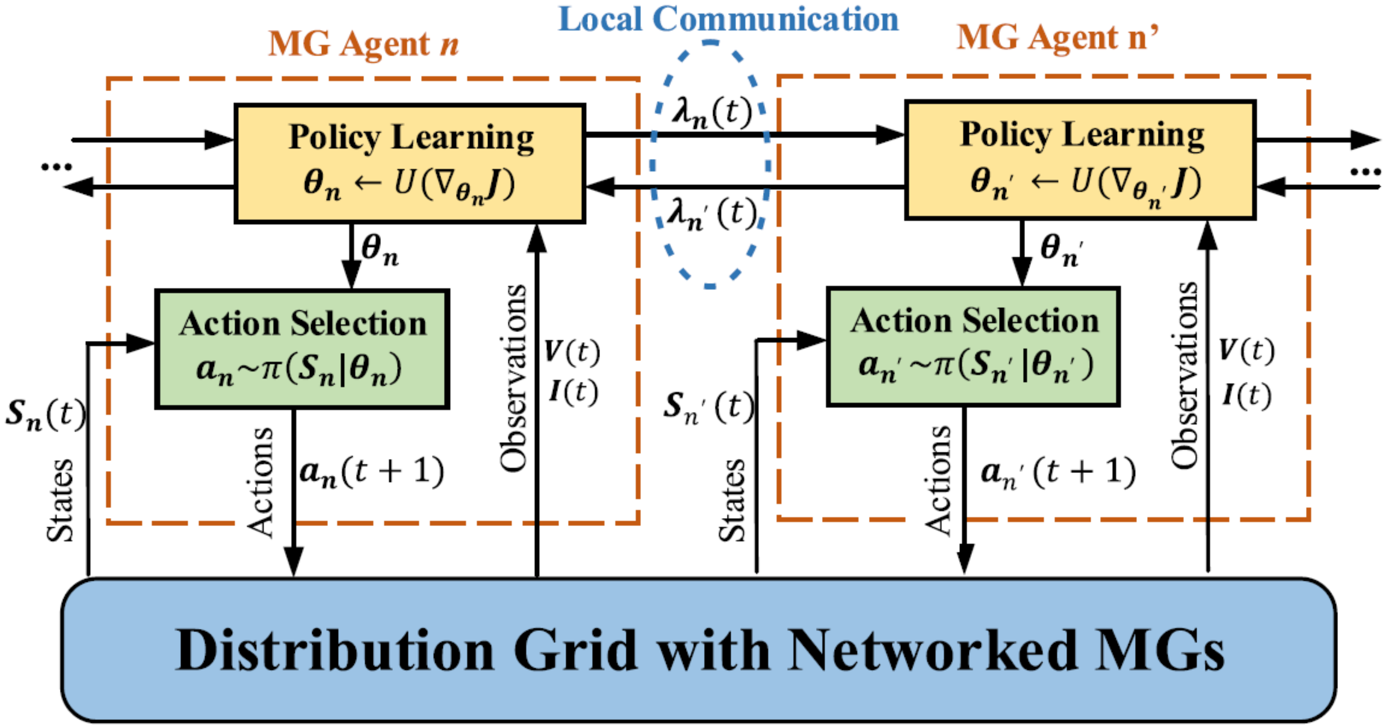
MASDRL-based Energy Management of MGs



Multi-Agent Safe Deep RL-based method:

- The energy management problem of networked MGs is reformulated as a policy learning problem with deep neural networks (DNNs).
- A **constrained gradient-based** training method is proposed that exploits the gradient information of the constraints and objective w.r.t. control actions and DNNs' learning parameters.
- A **distributed consensus-based** training process is proposed to decompose the training task among MG agents.

Q. Zhang, K. Dehghanpour, Z. Wang, F. Qiu and D. Zhao, "Multi-Agent Safe Policy Learning for Power Management of Networked Microgrids," in IEEE Transactions on Smart Grid, vol. 12, no. 2, pp. 1048-1062, March 2021.

MASDRL-based Energy Management of MGs

- **Safe RL**: **C**onstrained **M**arkov **D**ecision **P**rocess (CMDP) and trust region policy optimization method
- **Multi-agent RL**: **Scalability** and maintain **privacy** of MGs

Transfer | Approximate | Decompose

Energy management of networked MGs

$$\max_{x,u} \sum_{t=1}^{T} f(x(t), u(t))$$

$$\text{s.t} \quad g(x(t), u(t)) = 0$$

$$h(x(t), u(t)) \leq 0$$

$$LB \leq x(t) \leq UB$$

$$x \in \mathbb{R}, u \in \mathbb{I}$$

Safe policy learning

$$\pi^{t+1} = arg \max_{\pi_1,\ldots\pi_n} \sum_{n=1}^{N} J_{R_n}(\pi_n)$$

$$s.t. \quad a_n \sim \pi_n(S_n)$$

$$J_{C_m}(\pi) \leq d_m, \forall m$$

$$\Delta(\pi_n, \pi_n^t) \leq \delta, \forall n$$

- Control policy with DNNs

$$a_n \sim \pi_n(a_n|\theta_n) = \frac{1}{\sqrt{|\Sigma_n|(2\pi)^{D_n}}} e^{-\frac{1}{2}(a_n-\mu_n)^T \Sigma_n^{-1}(a_n-\mu_n)}$$

- Weights and bias of DNNs θ_{μ_n} and θ_{Σ_n}

$$\mu_n = DNN(S_n|\theta_{\mu_n})$$

$$\Sigma_n = DNN(S_n|\theta_{\Sigma_n})$$

Trust region policy optimization

$$\theta^{t+1} = arg \max_{\theta_1,\ldots\theta_n} \sum_{n=1}^{N} g_n^T(\theta_n - \theta_n^t)$$

$$s.t. \quad J_{C_m}(\theta^t) + b_m^T(\theta - \theta^t) \leq d_m, \forall m$$

$$\frac{1}{2}(\theta_n - \theta_n^t)^T H_n(\theta_n - \theta_n^t) \leq \delta, \forall n$$

- Gradient information

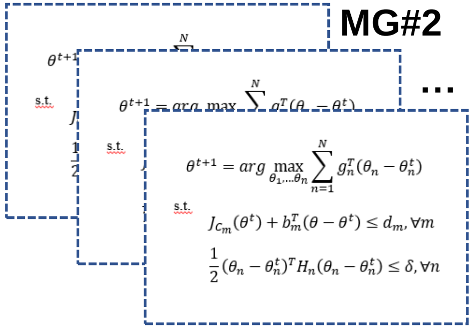
$$g_n = \nabla_\theta J_{R_n}$$

$$b_m = \nabla_\theta J_{C_m}$$

- KL divergence function and Fisher information matrix H_n

MG#1

MG#2



…

Primal-dual distributed method

- Primal update (global constraints)

$$\bar{\theta}_n(k) = \theta_n^t(k) - \rho_1\left(g_n\theta_n^t(k) + b_{m'}\theta_n^t(k)\bar{\lambda}_n(k)\right)$$

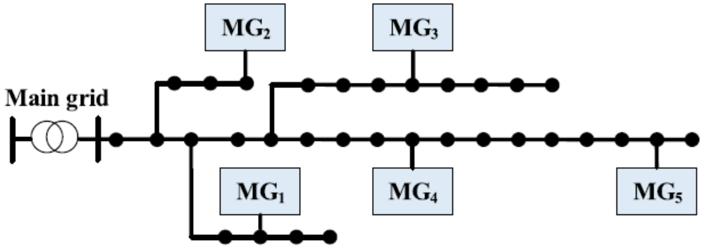
- Dual update (local constraints)

$$\lambda_n(k) = \left[\bar{\lambda}_n(k) + \rho_1\left(b_{m'}\theta_n^t(k+1) - d_{m'}\right)\right]^+$$

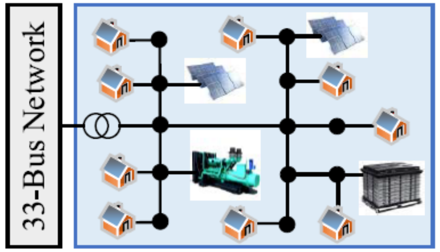
Q. Zhang, K. Dehghanpour, Z. Wang, F. Qiu and D. Zhao, "Multi-Agent Safe Policy Learning for Power Management of Networked Microgrids," in IEEE Transactions on Smart Grid, vol. 12, no. 2, pp. 1048-1062, March 2021.

MASDRL-based Energy Management of MGs

- Test distribution system with networked MGs



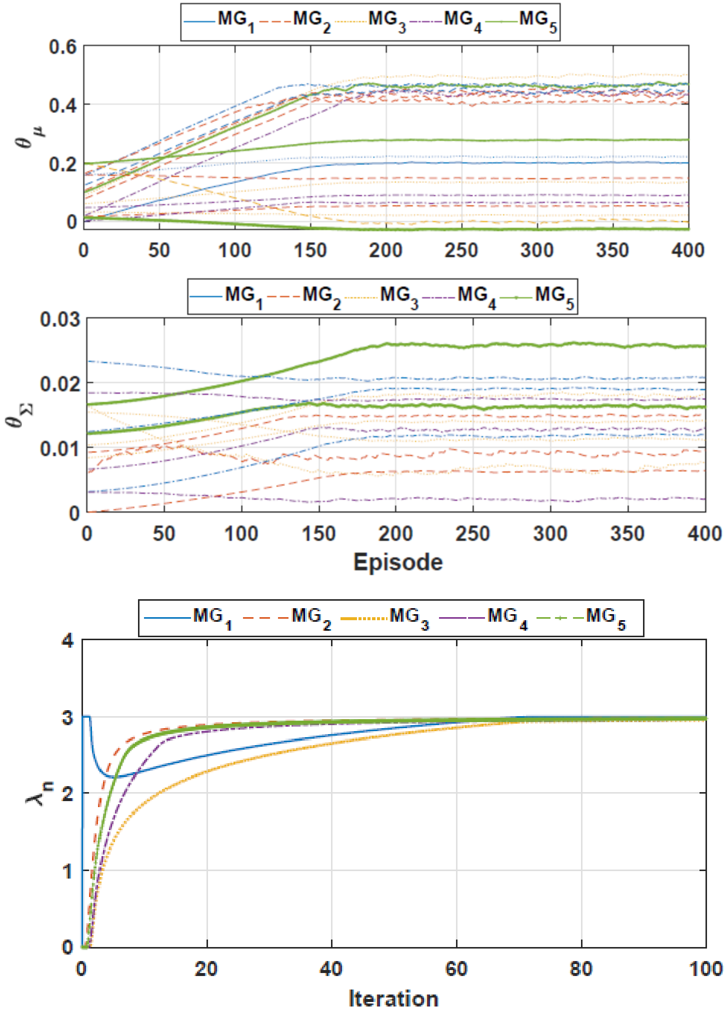
(a) 33-bus system for distribution network



(b) 13-bus system for MGs

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

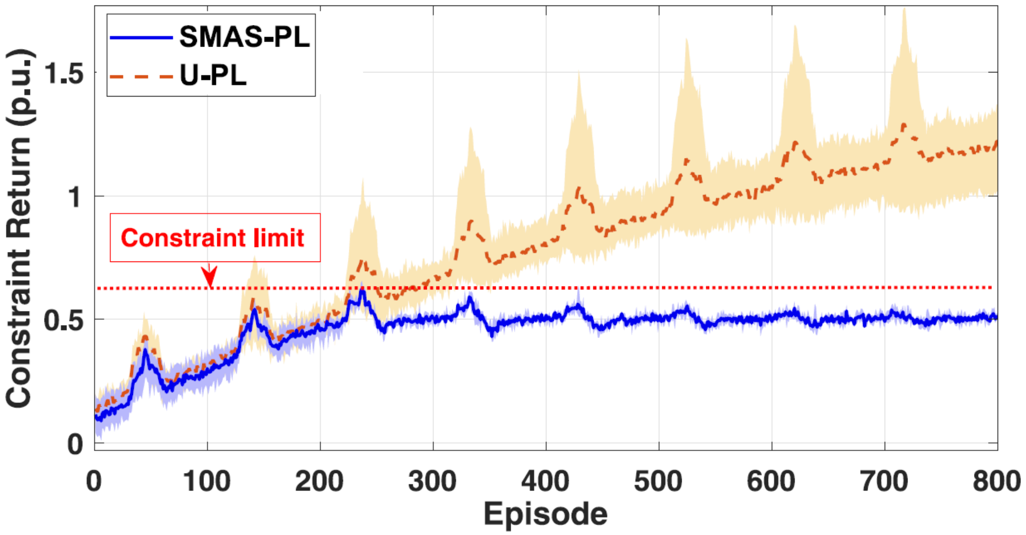
- Training results



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

	Cen. solver	DQN	SMAS-PL
Average daily cost ($)	1356.60	1928.4	1372.11
Average time (second)	145.50	10.30	1.40 (per agent)
MG privacy maintenance	No	No	Yes

- Constraint satisfaction



Q. Zhang, K. Dehghanpour, Z. Wang, F. Qiu and D. Zhao, "Multi-Agent Safe Policy Learning for Power Management of Networked Microgrids," in IEEE Transactions on Smart Grid, vol. 12, no. 2, pp. 1048-1062, March 2021.