Optimization and Reinforcement Learning for Multi-Agent Systems with Applications in Cyber Physical Networks

Yujie Tang

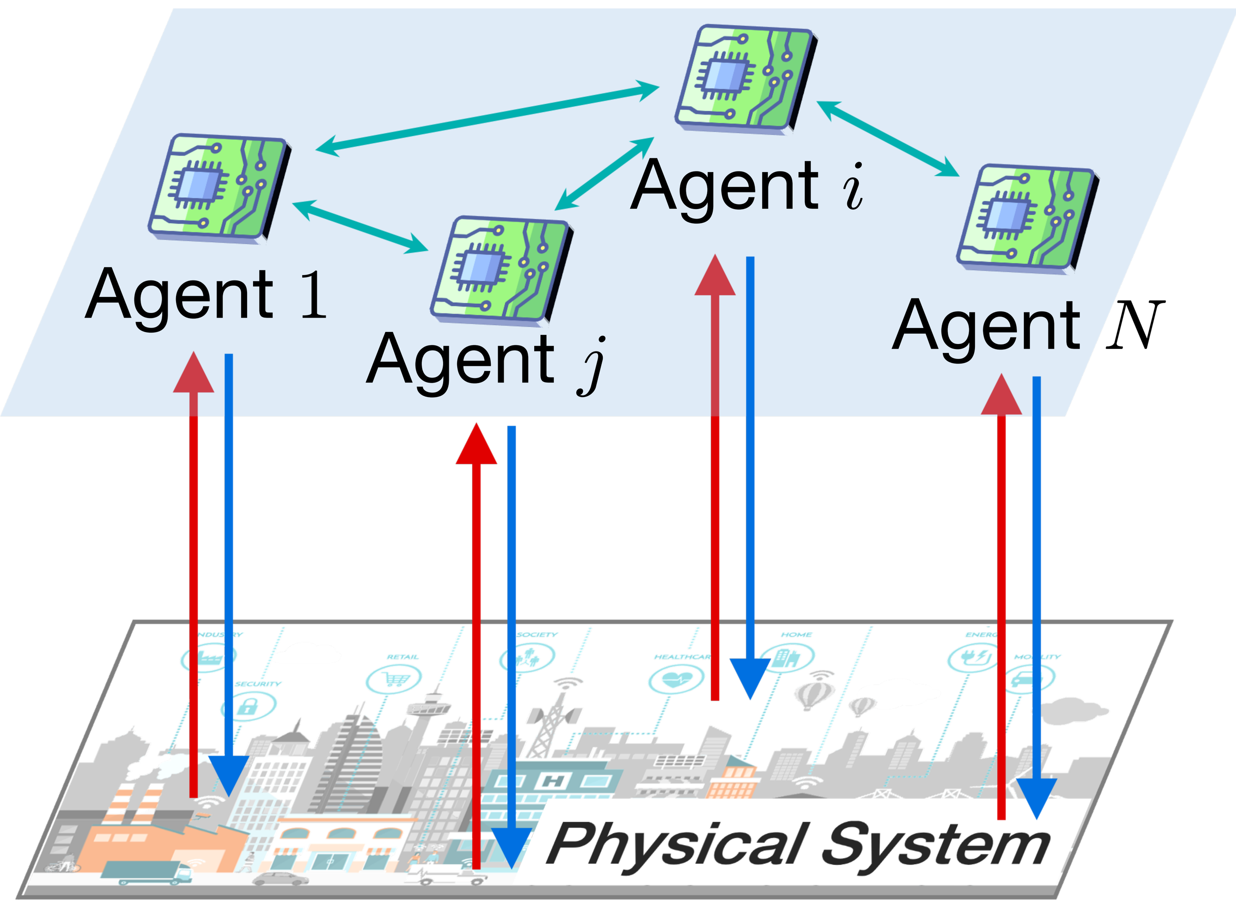
Harvard John A. Paulson School of Engineering and Applied Sciences

yujietang@seas.harvard.edu

scholar.harvard.edu/yujietang

Advisors: Na Li (Postdoc), Steven Low (PhD)

Overview



Observation

Action

Opportunities:

- Abundant, real-time data
- Computing power

Challenges:

- Unknown system model
- Partial observability
- Time-varying disturbances
- Rigorous performance guarantees
- …

How to exploit the opportunities to address the challenges?

Model-Free Multi-Agent Optimization and Reinforcement Learning

Zeroth-order optimization

Tailored to match specific observation & communication restrictions

Decentralized coordination

Static System: Optimization

- **Nonconvex consensus optimization**
  [Tang, Zhang and Li, TCNS 2020]
  - ✓ Convergence rate analysis
- **Social welfare optimization for multi-agent games**
  [Tang, Ren and Li, CDC & arXiv 2020]
  - ✓ Iteration/sample complexity analysis (convex & nonconvex)

Dynamical System: Control/RL

- **Distributed reinforcement learning for decentralized LQ control**
  [Li, Tang, Zhang and Li, TAC 2021]
  - Infinite-horizon average cost
  - Gaussian process noise
  - ✓ Stability guarantee
  - ✓ Sample complexity bound

Optimization Landscape of LQG

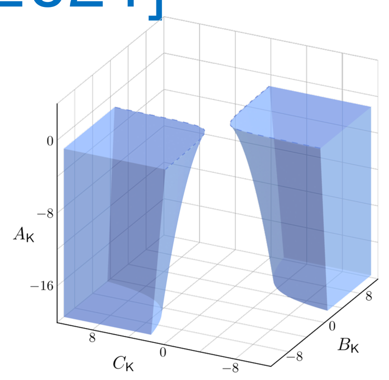
$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$
$$\text{s.t.} \quad \mathsf{K} \in \mathcal{C}$$

- LQG cost
- Set of full-order stabilizing *dynamic* controllers

[Tang, Zheng and Li, L4DC & arXiv 2021]

Domain $\mathcal{C}$

- May have one or two connected components.

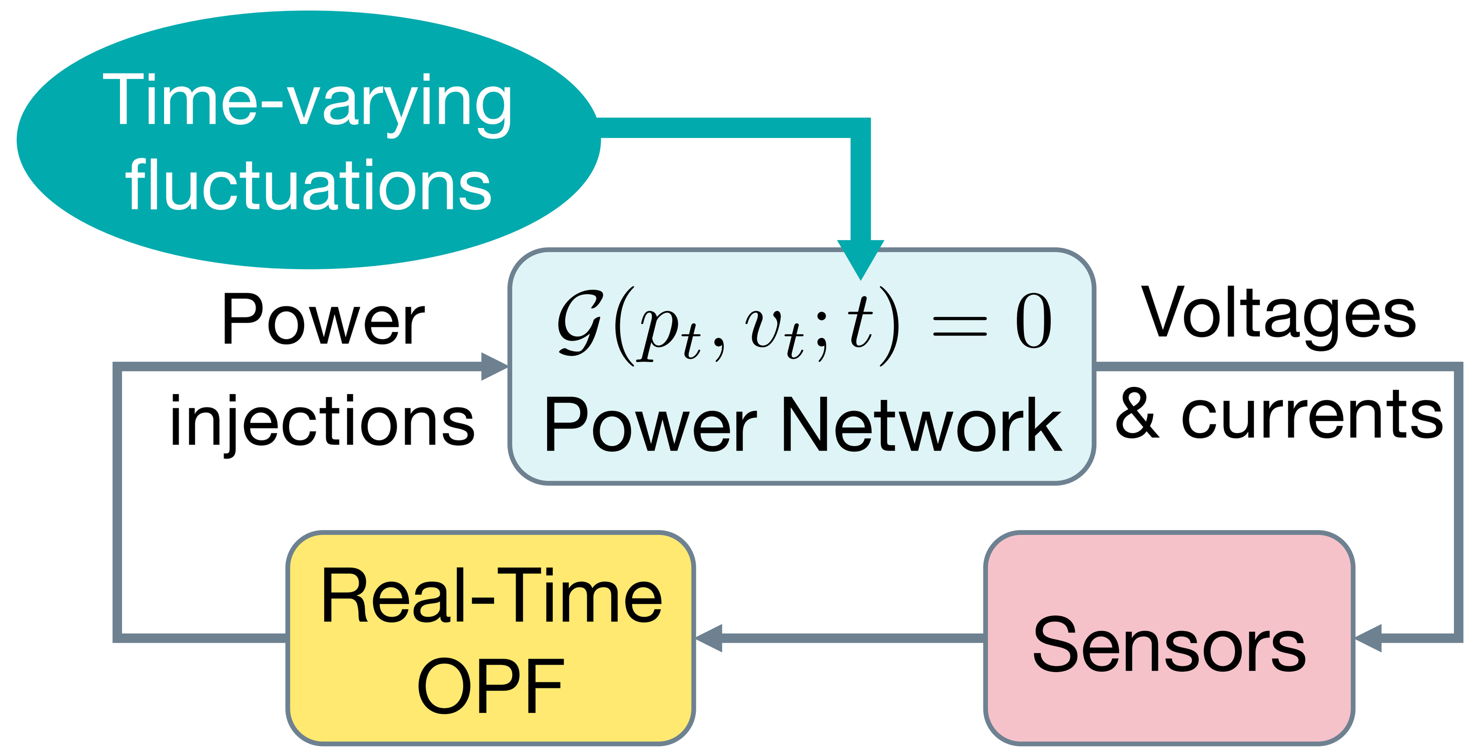


Objective $J(\mathsf{K})$

- Spurious stationary points and non-strict saddle points may exist.
- A stationary point is globally optimal if it is controllable & observable.

Real-Time Optimal Power Flow

[Tang, Dvijotham and Low, TSG 2017]



- Use the power network to solve PF equations.
- Update the problem data simultaneously with the optimization iterations.