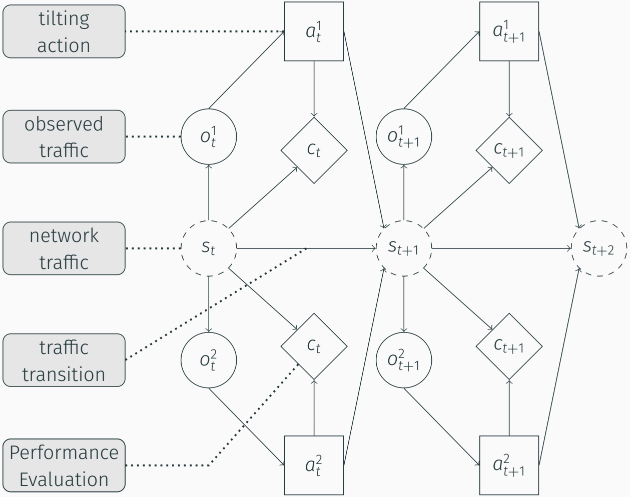
Digital-Twin-Enabled Predictive Traffic Sensing via Multi-Agent Risk-Constrained Online Learning

Tao Li taoli@nyu.edu
Department of Electrical and Computer Engineering
New York University

Apr. 17, 2025

The 2025 Annual East Coast Optimization Meeting
Center for Mathematics and Artificial Intelligence
George Mason University

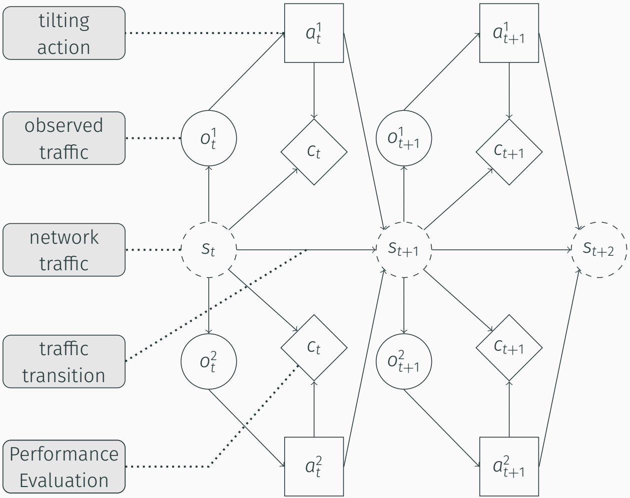
Pan-Tilt-Cameras Collaborative Sensing



DT-Enabled Predictive Sensing

Modeling

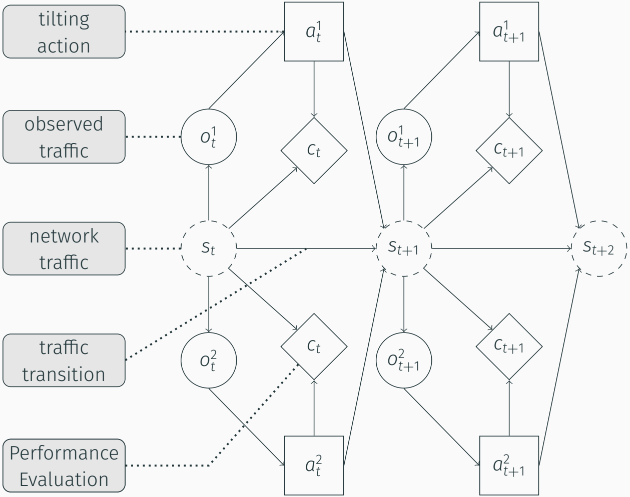
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;



DT-Enabled Predictive Sensing

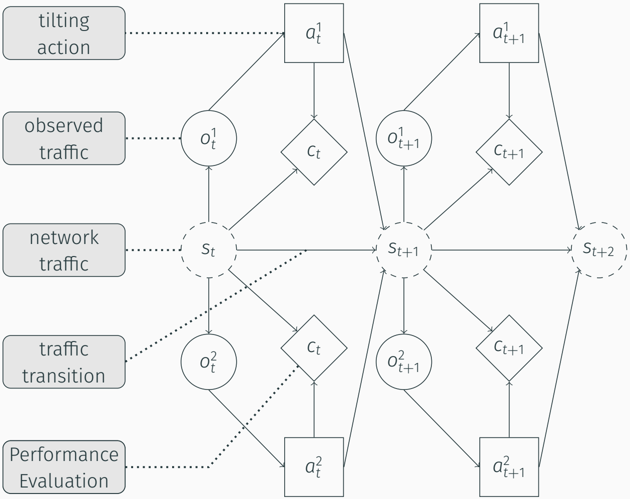
Modeling

- Env: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- Action: $a_t^i \in \{0, 1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;



Modeling

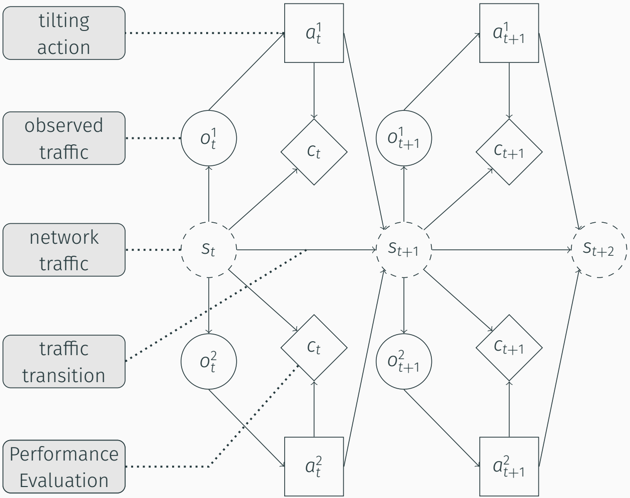
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0, 1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;



DT-Enabled Predictive Sensing

Modeling

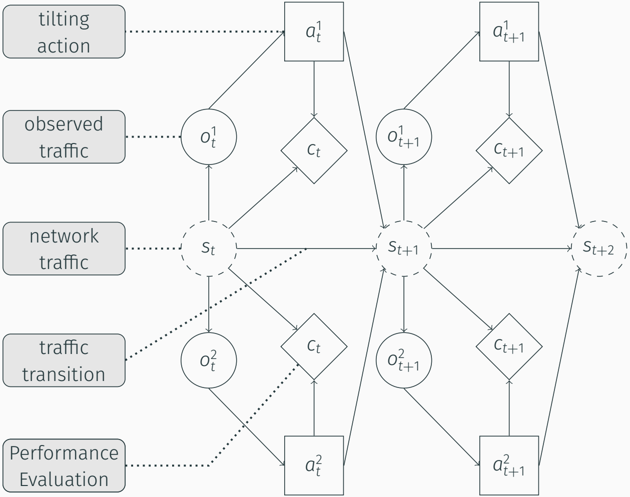
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0,1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;
    - ⚠ notoriously hard to estimate [Yuan et al., Anal. Methods Accid. Res'22];



DT-Enabled Predictive Sensing

Modeling

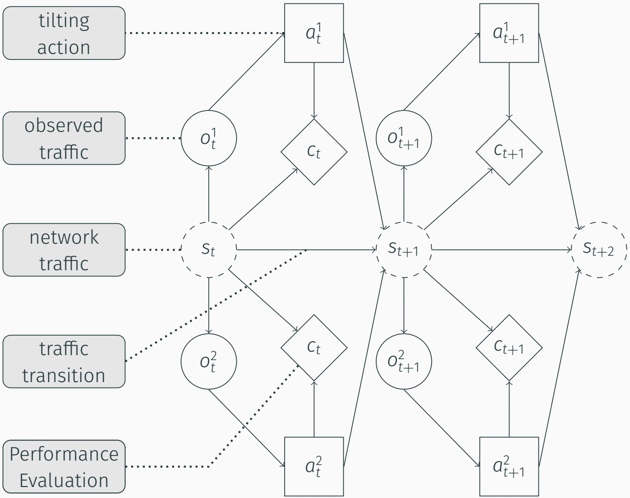
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0,1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;
  - ⚠ notoriously hard to estimate [Yuan et al., Anal. Methods Accid. Res'22];
  - 🖥 DT can simulate r_t from s_t;



DT-Enabled Predictive Sensing

Modeling

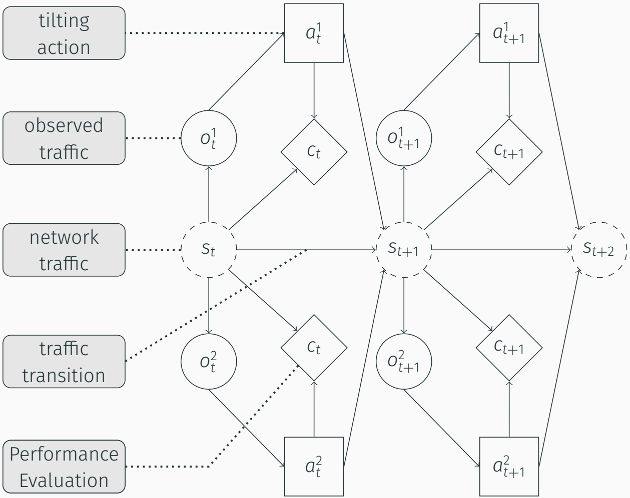
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0,1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;
  - ⚠ notoriously hard to estimate [Yuan et al., Anal. Methods Accid. Res'22];
  - DT can simulate r_t from s_t;
- **Performance**: joint action $a_t = \vee_{i \in N} a_t^i$



DT-Enabled Predictive Sensing

Modeling

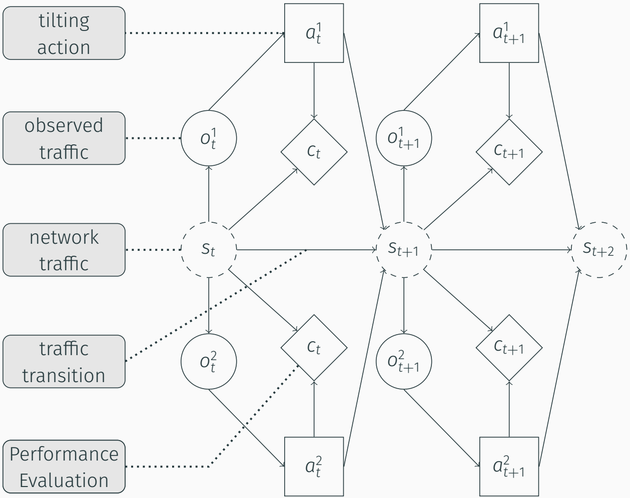
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0,1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;
  - ⚠ notoriously hard to estimate [Yuan et al., Anal. Methods Accid. Res'22];
  - DT can simulate r_t from s_t;
- **Performance**: joint action $a_t = \vee_{i \in N} a_t^i$
  - Information Acquisition: $f(a_t; s_t) = \langle a_t, s_t \rangle$;



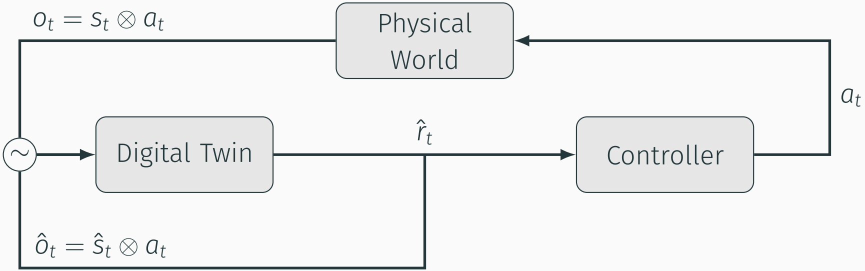
DT-Enabled Predictive Sensing

Modeling

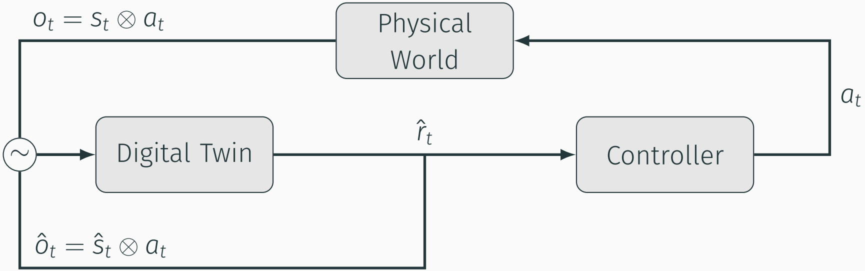
- **Env**: $\mathcal{G} = \langle N, E \rangle$, $s_t \in \mathbb{R}^E$, $\{s_t\}_{t=1}^T$ adversarially picked by Nature;
- **Action**: $a_t^i \in \{0,1\}^E$, $o_t^i = s_t \otimes a_{t-1}^i$;
- **Latent Variable**: safety risk $r_t \in \mathbb{R}^E$, $(s_t, r_t) \sim \mathbb{P}$;
  - ⚠ notoriously hard to estimate [Yuan et al., Anal. Methods Accid. Res'22];
  - DT can simulate r_t from s_t;
- **Performance**: joint action $a_t = \vee_{i \in N} a_t^i$
  - Information Acquisition: $f(a_t; s_t) = \langle a_t, s_t \rangle$;
  - Predictive Monitoring: $g(a_t; r_t) = \langle a_t, r_t \rangle$.



The Challenge of Dual Control



The Challenge of Dual Control



Balancing the Dual Effect

- Directing: each control directs preemptive sensing $\max g(a_t; \hat{r}_t)$;
- Probing: each control affects the perception of uncertainty $\hat{r}_t$ via traffic synchronization $\max f(a_t; s_t)$

$$\max_{\pi_t} f(\pi_t; s_t) \quad \text{s.t.} \quad g(\pi_t; \hat{r}_t) \geq \beta .$$

randomized (π_t) — Hyperparameter (β)

Correlated Lagrangian for Multi-Agent Constrained Online Learning

$$\min_{\pi_t} -f(\pi_t; s_t) \quad \text{s.t.} \quad h(\pi_t; \hat{r}_t) \triangleq \beta - g(\pi_t; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \ldots, T\}$$

Correlated Lagrangian for Multi-Agent Constrained Online Learning

$$\min_{\pi_t} -f(\pi_t; s_t) \quad \text{s.t.} \quad h(\pi_t; \hat{r}_t) \triangleq \beta - g(\pi_t; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}$$

Online Learning

Time-varying objectives and constraints without prior information: find $\pi_t \leftarrow Alg(\pi_{1:t-1}, o_{1:t-1})$
No-Regret: $\max_{\pi \in \Pi} \sum_{t=1}^{T} f(\pi; s_t) - \sum_{t=1}^{T} f(\pi_t; s_t) \sim o(T), \quad \sum_{t=1}^{T} h(\pi_t; \hat{r}_t) \sim o(T).$

Correlated Lagrangian for Multi-Agent Constrained Online Learning

$$\min_{\pi_t} -f(\pi_t; s_t) \quad \text{s.t.} \quad h(\pi_t; \hat{r}_t) \triangleq \beta - g(\pi_t; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}$$

Online Learning

Time-varying objectives and constraints without prior information: find $\pi_t \leftarrow Alg(\pi_{1:t-1}, o_{1:t-1})$

No-Regret: $\max_{\pi \in \Pi} \sum_{t=1}^{T} f(\pi; s_t) - \sum_{t=1}^{T} f(\pi_t; s_t) \sim o(T), \quad \sum_{t=1}^{T} h(\pi_t; \hat{r}_t) \sim o(T)$.

Lagrangian Multiplier and Primal-Dual

KKT condition and saddle point: $\min_{\pi \in \Pi} \max_{\lambda \in \mathbb{R}_+} \mathcal{L}(\pi, \lambda; s_t, \hat{r}_t) \triangleq -f(\pi; s_t) + \lambda h(\pi; \hat{r}_t)$.

$$\pi_{t+1} = \text{Proj}_{\Pi}[\pi_t - \gamma_t \nabla_\pi \mathcal{L}(\pi_t, \lambda_t; s_t, \hat{r}_t)], \quad \lambda_{t+1} = \text{Proj}_{R_+}[\lambda_t + \gamma_t \nabla_\lambda \mathcal{L}(\pi_t, \lambda_t; s_t, \hat{r}_t)].$$

Correlated Lagrangian for Multi-Agent Constrained Online Learning

$$\min_{\pi_t} -f(\pi_t; s_t) \quad \text{s.t.} \quad h(\pi_t; \hat{r}_t) \triangleq \beta - g(\pi_t; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \ldots, T\}$$

Online Learning

Time-varying objectives and constraints without prior information: find $\pi_t \leftarrow Alg(\pi_{1:t-1}, o_{1:t-1})$

No-Regret: $\max_{\pi \in \Pi} \sum_{t=1}^{T} f(\pi; s_t) - \sum_{t=1}^{T} f(\pi_t; s_t) \sim o(T), \quad \sum_{t=1}^{T} h(\pi_t; \hat{r}_t) \sim o(T)$.

Lagrangian Multiplier and Primal-Dual

KKT condition and saddle point: $\min_{\pi \in \Pi} \max_{\lambda \in \mathbb{R}_+} \mathcal{L}(\pi, \lambda; s_t, \hat{r}_t) \triangleq -f(\pi; s_t) + \lambda h(\pi; \hat{r}_t)$.

$$\pi_{t+1} = \text{Proj}_{\Pi}[\pi_t - \gamma_t \nabla_\pi \mathcal{L}(\pi_t, \lambda_t; s_t, \hat{r}_t)], \quad \lambda_{t+1} = \text{Proj}_{R_+}[\lambda_t + \gamma_t \nabla_\lambda \mathcal{L}(\pi_t, \lambda_t; s_t, \hat{r}_t)].$$

Correlated Lagrangian Multiplier in Multi-Agent Learning

NYC (with over 15,000 cameras) calls for distributed control:

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \ldots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \text{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \text{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \mathrm{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \mathrm{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \text{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \text{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica'00], [Liu-Li-Zhu, TISPN'23]

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \text{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \text{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica'00], [Liu-Li-Zhu, TISPN'23]
- What does the asymptotics look like?

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i,\pi_t^{-i}} -f(\pi_t^i,\pi_t^{-i};s_t) \quad \text{s.t.} \quad h(\pi_t^i,\pi_t^{-i};\hat{r}_t) \leq 0, \quad t \in \{1,2,\dots,T\}.$$

$$\min_{\pi^i\in\Pi^i} \max_{\lambda^i\in\mathbb{R}_+} \mathcal{L}^i(\pi^i,\pi_t^{-i},\lambda^i;s_t,\hat{r}_t) \triangleq -f(\pi^i,\pi_t^{-i};s_t) + \lambda^i h(\pi^i,\pi_t^{-i};\hat{r}_t).$$

$$\pi_{t+1}^i = \mathrm{Proj}_\Pi[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \mathrm{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica'00], [Liu-Li-Zhu, TISPN'23]
- What does the asymptotics look like?

- **Conjectural online learning with first-order beliefs in asymmetric information stochastic games**, *IEEE Conference on Decision and Control*, 2024.

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \ldots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \mathrm{Proj}_\Pi[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \mathrm{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica'00], [Liu-Li-Zhu, TISPN'23]
- What does the asymptotics look like?

- **Conjectural online learning with first-order beliefs in asymmetric information stochastic games**, *IEEE Conference on Decision and Control*, 2024.
- **Adaptive Security Response Strategies through Conjectural Online Learning**, IEEE Transactions on Information Forensics and Security, 2025

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \dots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \mathrm{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \mathrm{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica'17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.'15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica'00], [Liu-Li-Zhu, TISPN'23]
- What does the asymptotics look like?

- **Conjectural online learning with first-order beliefs in asymmetric information stochastic games**, *IEEE Conference on Decision and Control*, 2024.
- **Adaptive Security Response Strategies through Conjectural Online Learning**, IEEE Transactions on Information Forensics and Security, 2025
- **Multi-level traffic-responsive tilt camera surveillance through predictive correlated online learning**, *Transportation Research Part C: Emerging Technologies*, 2024.

Takeaway Message

Multi-Agent Constrained Correlated Learning

$$\min_{\pi_t^i, \pi_t^{-i}} -f(\pi_t^i, \pi_t^{-i}; s_t) \quad \text{s.t.} \quad h(\pi_t^i, \pi_t^{-i}; \hat{r}_t) \leq 0, \quad t \in \{1, 2, \ldots, T\}.$$

$$\min_{\pi^i \in \Pi^i} \max_{\lambda^i \in \mathbb{R}_+} \mathcal{L}^i(\pi^i, \pi_t^{-i}, \lambda^i; s_t, \hat{r}_t) \triangleq -f(\pi^i, \pi_t^{-i}; s_t) + \lambda^i h(\pi^i, \pi_t^{-i}; \hat{r}_t).$$

$$\pi_{t+1}^i = \text{Proj}_{\Pi}[\pi_t^i - \gamma_t \nabla_i \mathcal{L}^i], \quad \lambda_{t+1}^i = \text{Proj}_{R_+}[\lambda_t^i + \gamma_t \nabla_\lambda \mathcal{L}^i].$$

- Online constraints balance the directing and probing effect; ⚠the problem is too model-free.
- ✗Adaptive control [Heirung-Ydstie-Foss, Automatica’17]; ✗Bandit learning with partial monitoring [Perchet-Quincampoix, Math. Oper. Res.’15];
- Correlated learning in game theory [Hart-Mas-Colell, Econometrica’00], [Liu-Li-Zhu, TISPN’23]
- What does the asymptotics look like?

- **Conjectural online learning with first-order beliefs in asymmetric information stochastic games**, *IEEE Conference on Decision and Control*, 2024.
- **Adaptive Security Response Strategies through Conjectural Online Learning**, IEEE Transactions on Information Forensics and Security, 2025
- **Multi-level traffic-responsive tilt camera surveillance through predictive correlated online learning**, *Transportation Research Part C: Emerging Technologies*, 2024.
-