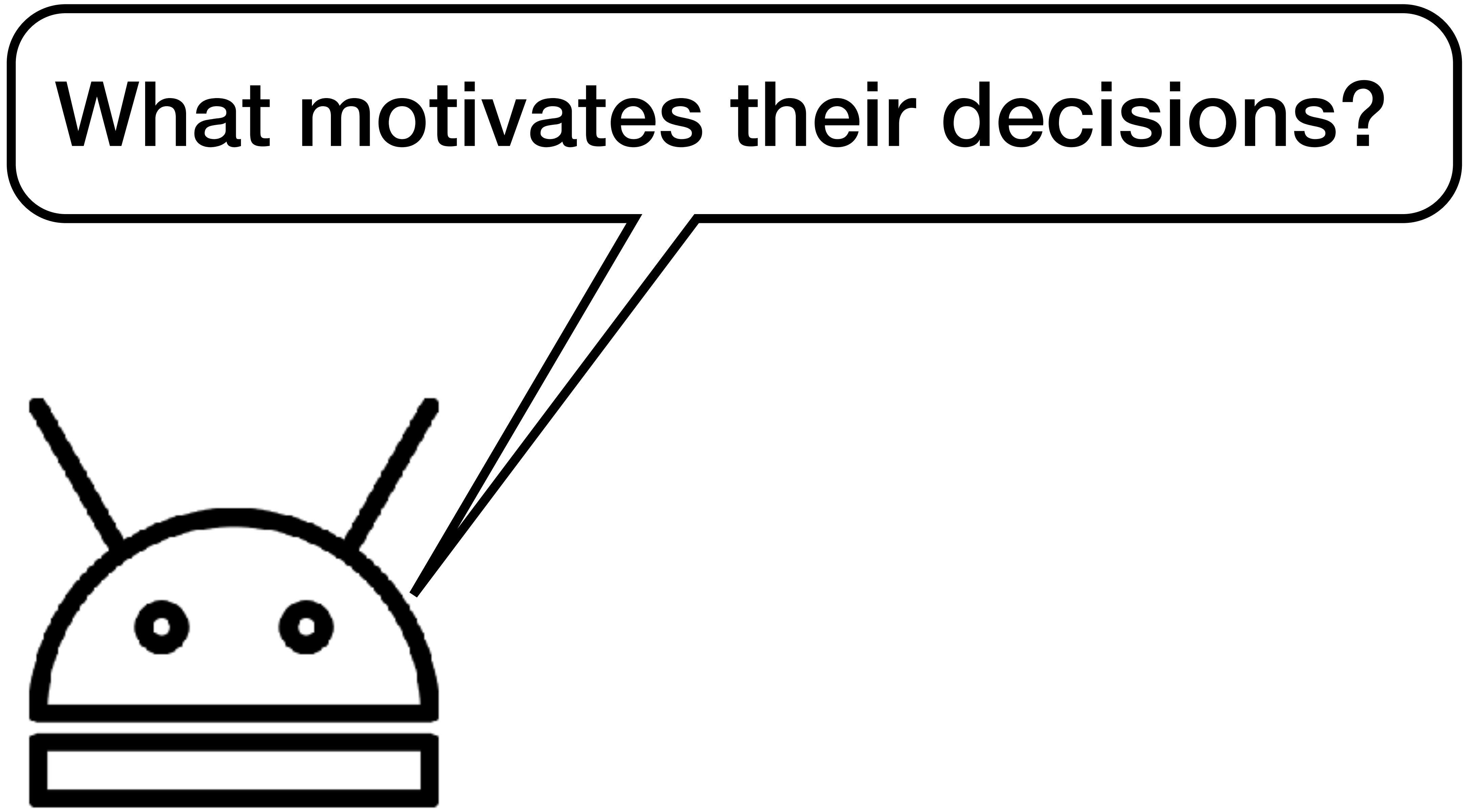
Inverse Games: Inferring Motives from Interactions

Yue Yu

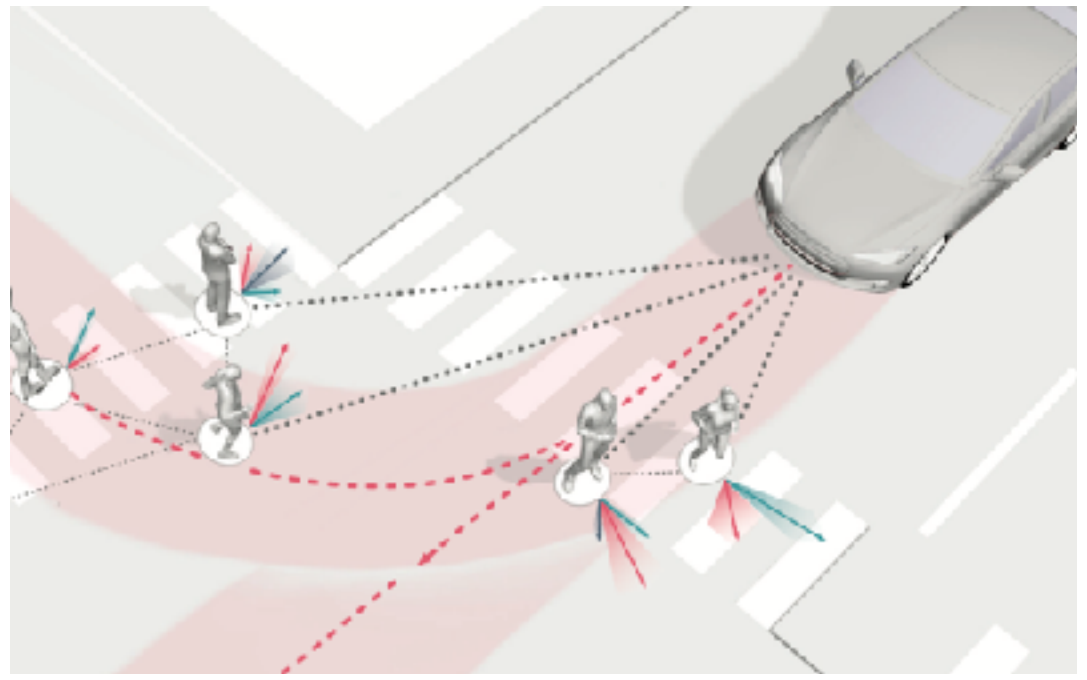
Collaborators:
Shenghui Chen, Jacob Levy, David Fridovich-Keil, Negar Mehr, Ufuk Topcu

TEXAS
The University of Texas at Austin

Inverse Games:
Explaining and Predicting Interactions



Service Provider Competition



Human-Robot Interactions

Multiagent Coordination

Resource Allocation

https://smallbusinessbonfire.com https://ai.stanford.edu/blog
https://www.collinsaerospace.com https://www.gsquaredcfo.com/blog

Nash Equilibrium in Discrete Games

cost due to player's own decision

Nash Equilibrium Condition

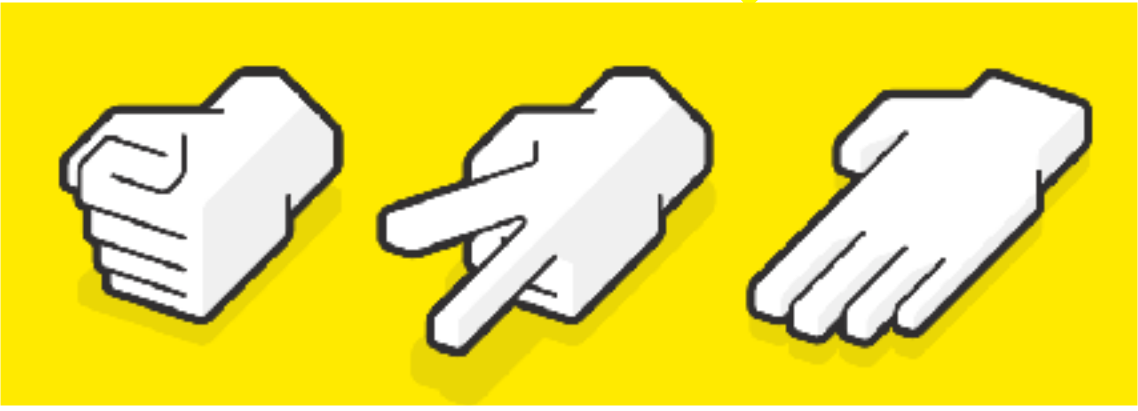
cost due to other players' decisions

$$x_i^\star \in \underset{x_i}{\operatorname{argmin}} \quad \frac{1}{2} x_i^\top C_{ii} x_i + b_i^\top x_i + \textstyle\sum_{j \neq i} x_i^\top C_{ij} x_j^\star$$

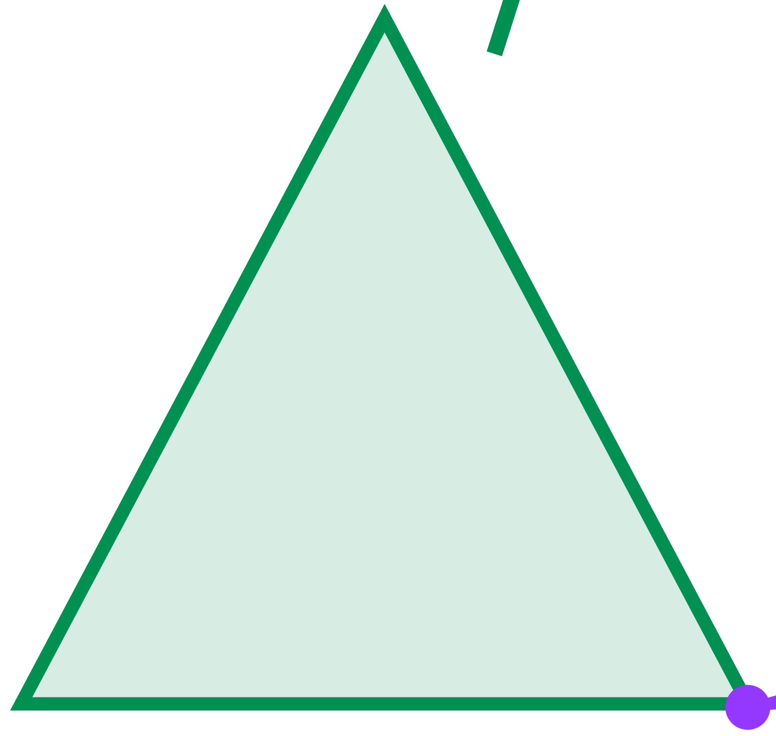
$$\text{s.t.} \quad x_i^\top \mathbf{1} = 1,\ x_i \geq 0$$

Only explains rational decisions!

30% 50% 20%



EsquireME



Always play Rock!

Bounded Rationality: Modeling Noisy Decision-Making

Quantal Response Equilibrium Condition

$$x_i^\star \in \underset{x_i}{\operatorname{argmin}} \quad \frac{1}{2} x_i^\top C_{ii} x_i + b_i^\top x_i + \sum_{j \neq i} x_i^\top C_{ij} x_j^\star \; + \; x_i^\top \ln(x_i)$$

$$\text{s.t.} \quad x_i^\top \mathbf{1} = 1,\; x_i \geq 0.$$

Entropy captures noisy behavior!

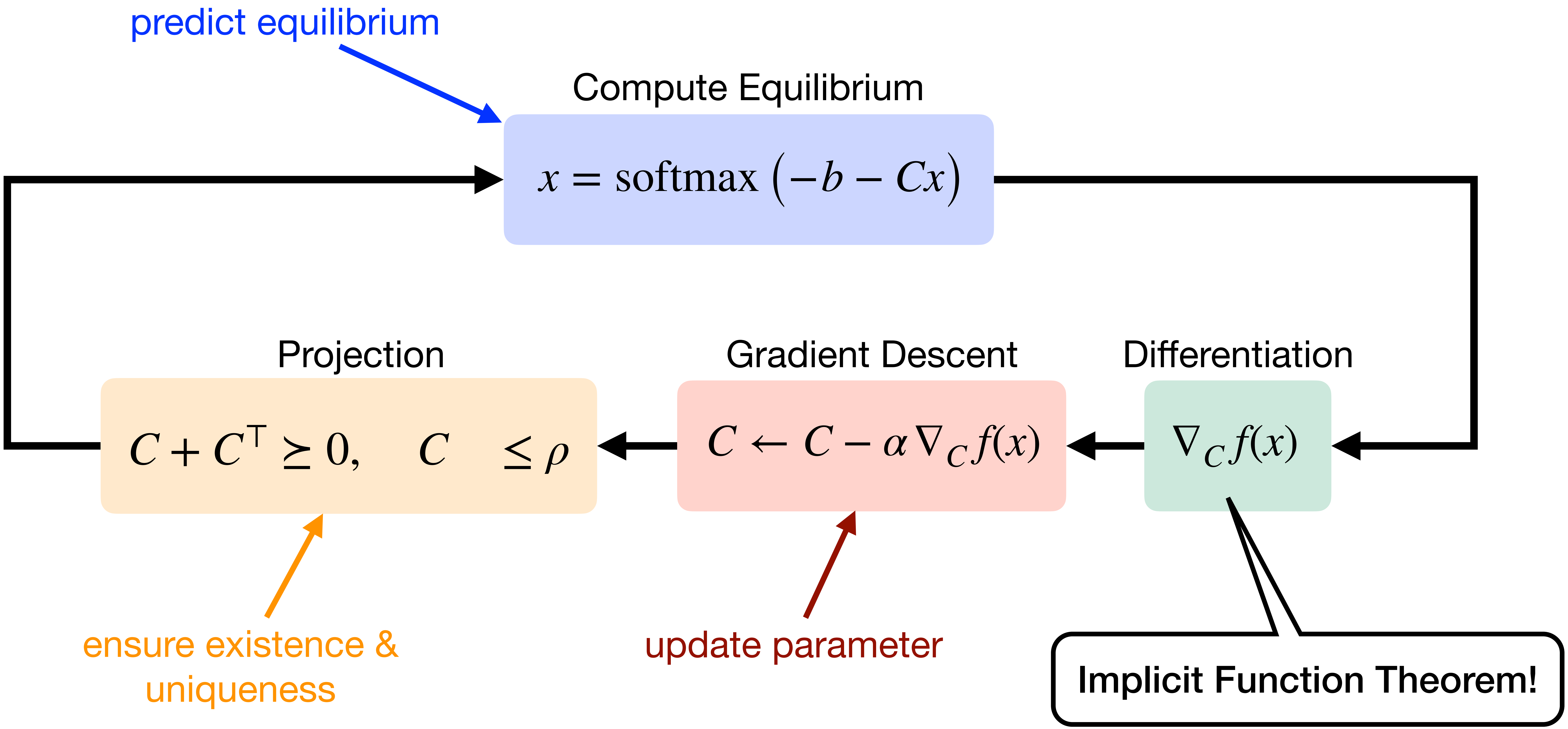
McKelvey & Palfrey ('95, '98)

⇕

Solve equations, predict equilibrium!

$$x^\star = \operatorname{softmax}\left(-b - Cx^\star\right)$$

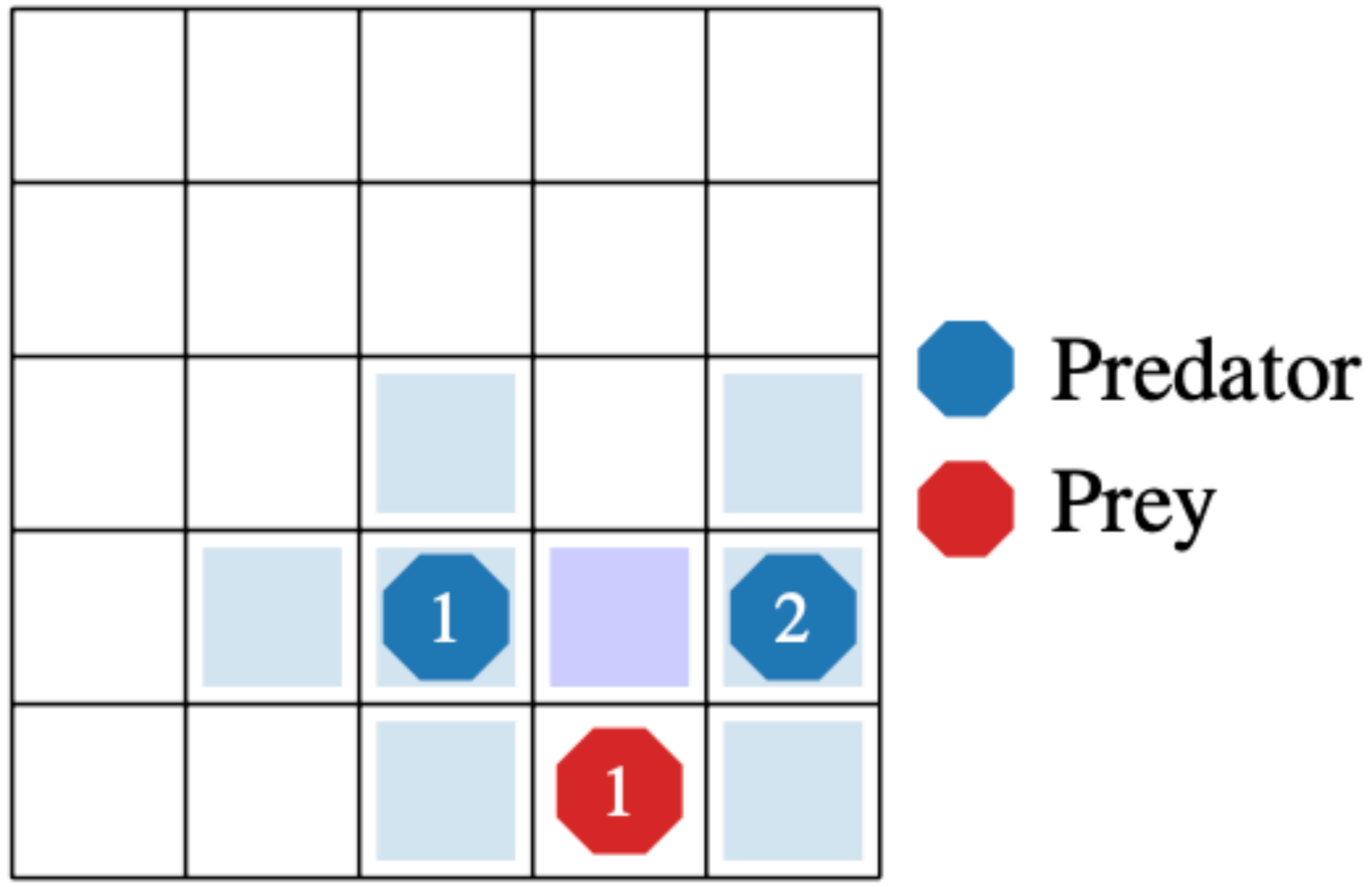
Learning Motives via Implicit Differentiation



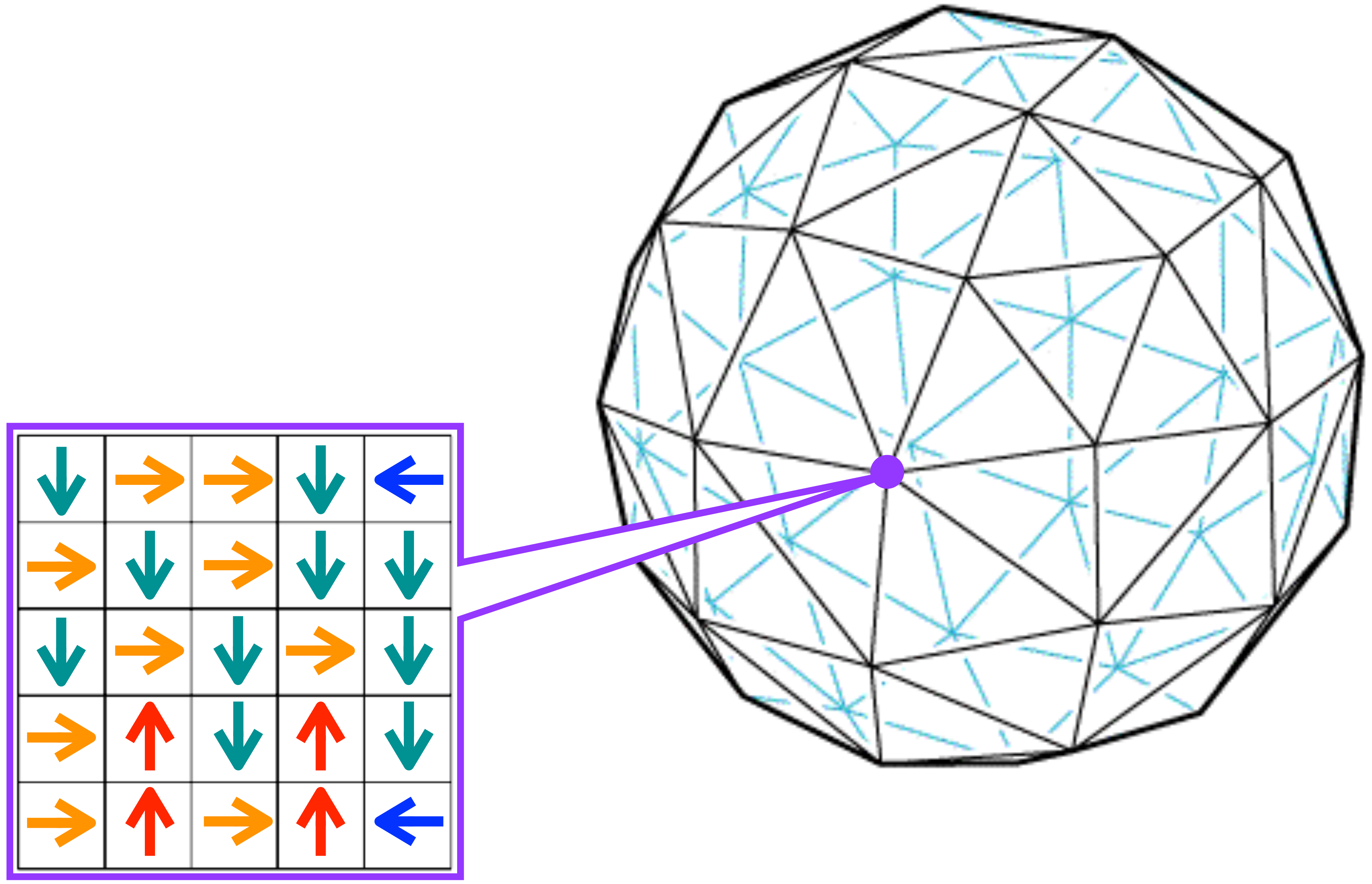
Differentiating through Optimization

Research Area	Optimization	Differentiation	Problem Dimension
Deep Learning	ReLU, Sigmoid, Softmax Amos('19)	Explicit function	Ridiculously high
Inverse Learning Bilevel Optimization	Convex Optimization Agrawal et al ('19)	Least-squares	High
Games Inverse Learning	Nonlinear Least-Squares Amos ('22), **Yu et al ('22)**	Least-squares	Medium (so far)

What about Games with Dynamic Decision-Making?



Multiplayer Markov Game



Policy Polytope

Soft-Bellman Equilibrium in Affine Markov Games

Equilibrium Conditions

$$Y^i_{sa} = \sum_{t=0}^{\infty} \gamma^t \mathbb{P}(S^i_t = s, A^i_t = a)$$

$$\Pi^i_{sa} = \mathbb{P}(A^i_t = a \quad S^i_t = s) = \frac{Y^i_{sa}}{\sum_{a=1}^{m} Y^i_{sa}}$$

Policies & state-action frequencies

$$Q^i_{sa} = R^i_{sa} + \gamma \sum_{s'=1}^{n} T^i_{sas'} v^i_{s'}$$

$$v^i_s = \log\left(\sum_{a=1}^{m} \exp(Q^i_{sa})\right)$$

$$\Pi^i_{sa} = \frac{\exp(Q^i_{sa})}{\sum_{a=1}^{m} \exp(Q^i_{sa})}$$

Soft-Bellman equations

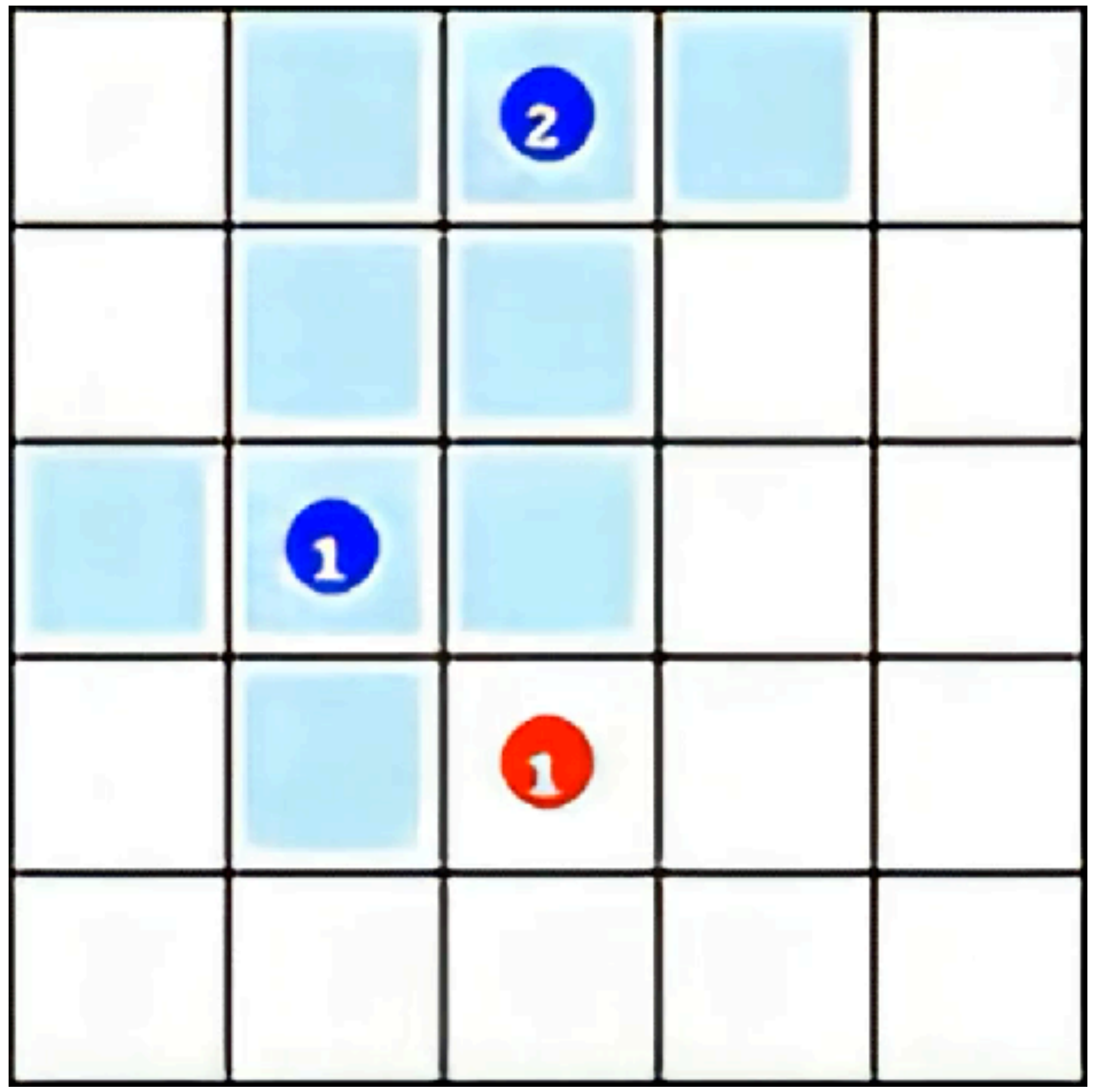
$$\textbf{vec}(R^i) = b^i + \sum_{j=1}^{p} C^{ij} \textbf{vec}(Y^j)$$

Affine cost coupling

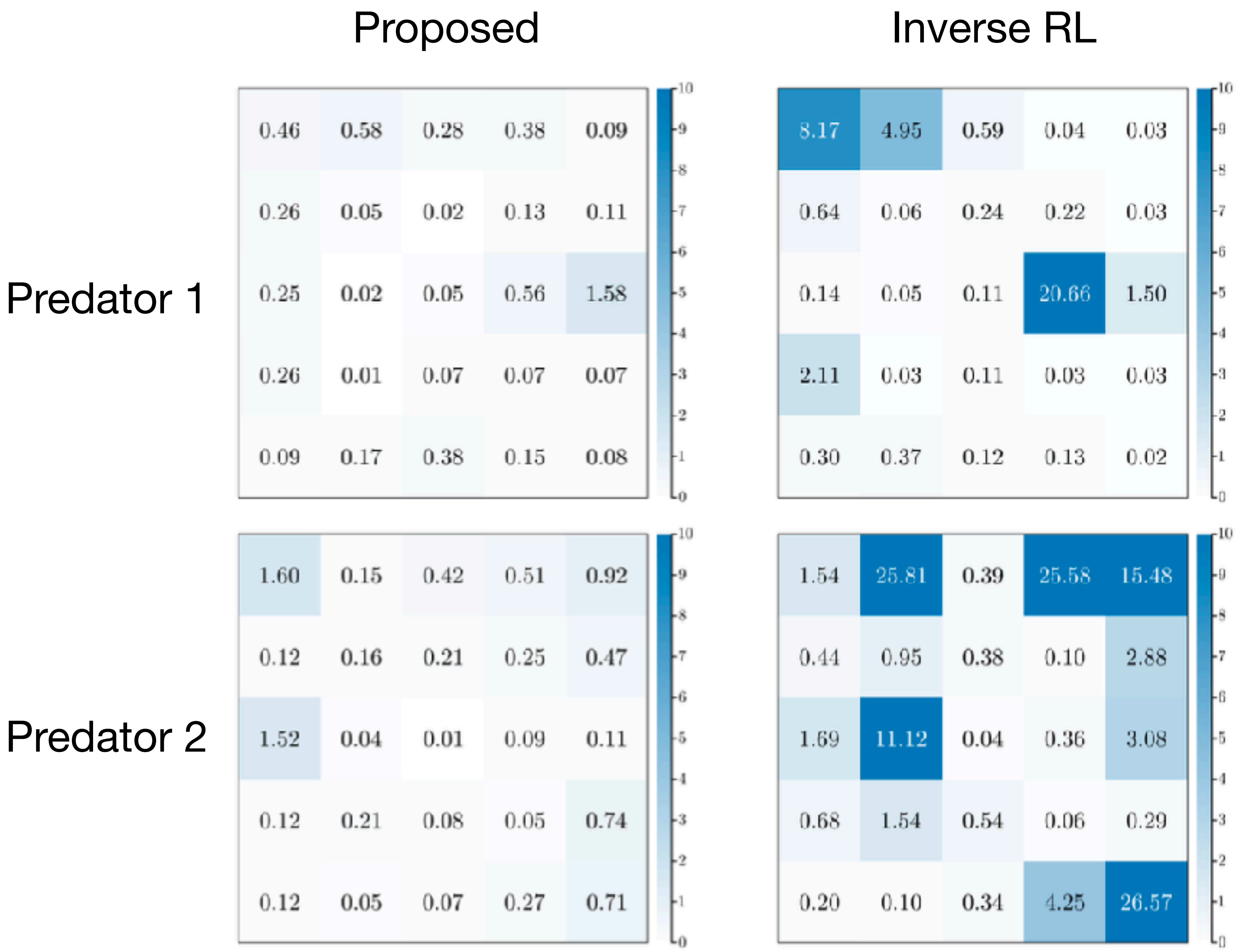
Same problem as before, just more equations!

(more complicated polytopes)

Soft-Bellman Equilibrium in Markov Games



A Three-Player Markov Game



KL-Divergence of Policies

How to Provoke Informative Actions in Games?

The Battle of Seven Potters

The Potters want to infer which Potter is the bad wizard chasing. What should the Potters do?

Stick Together

Split Up

Stackelberg Trajectory Games

Rational Leader

(Potters)

$$\underset{u^{\mathrm{L}}_{0:\tau-1},\, x^{\mathrm{L}}_{0:\tau}}{\text{minimize}} \quad \mathbb{E}[f(x^{\mathrm{L}}_{0:\tau}, x^{\mathrm{F}}_{0:\tau})] + g(u^{\mathrm{L}}_{0:\tau-1})$$

$$\text{subject to} \quad x^{\mathrm{L}}_{t+1} = A^{\mathrm{L}} x^{\mathrm{L}}_t + B^{\mathrm{L}} u^{\mathrm{L}}_t,\ u^{\mathrm{L}} \in \mathbb{U}$$

Leader knows Follower!

Boundedly Rational Follower

(Bad Wizard)

$$\underset{\mu_{0:\tau-1},\Sigma_{0:\tau-1}}{\text{minimize}} \quad \mathbb{E}\left[\sum_{t=0}^{\tau} \ x^{\mathrm{F}}_t - M x^{\mathrm{L}}_t \ {}^{2}_{Q^{\mathrm{F}}} + \sum_{t=0}^{\tau-1} \ u^{\mathrm{F}}_t \ {}^{2}_{R^{\mathrm{F}}}\right] + \sum_{t=0}^{\tau-1} \log\det \Sigma_t^{-1}$$

$$\text{subject to} \quad x^{\mathrm{F}}_{t+1} = A^{\mathrm{F}} x^{\mathrm{F}}_t + B^{\mathrm{F}} u^{\mathrm{F}}_t$$

$$u^{\mathrm{F}}_t \ \ x^{\mathrm{F}}_t \sim \mathcal{N}(\mu_t, \Sigma_t)$$

Causal entropy captures bounded rationality!

What If Leader Does Not Know Follower's Type?

Rational Leader

(Potters)

$$\underset{u^{L}_{0:\tau-1},\, x^{L}_{0:\tau}}{\text{minimize}} \quad \mathbb{E}[f(x^{L}_{0:\tau}, x^{F}_{0:\tau})] + g(u^{L}_{0:\tau-1})$$

$$\text{subject to} \quad x^{L}_{t+1} = A^{L}x^{L}_{t} + B^{L}u^{L}_{t},\ u^{L} \in \mathbb{U}$$

Leader only knows $M \in \{M^1, M^2, \ldots, M^d\}$

How to pinpoint Follower's type?

Boundedly Rational Follower

(Bad Wizard)

$$\underset{\mu_{0:\tau-1}, \Sigma_{0:\tau-1}}{\text{minimize}} \quad \mathbb{E}\left[\sum_{t=0}^{\tau} \left\| x^{F}_{t} - M x^{L}_{t} \right\|^2_{Q^{F}} + \sum_{t=0}^{\tau-1} \left\| u^{F}_{t} \right\|^2_{R^{F}}\right] + \sum_{t=0}^{\tau-1} \log\det \Sigma_t^{-1}$$

$$\text{subject to} \quad x^{F}_{t+1} = A^{F}x^{F}_{t} + B^{F}u^{F}_{t}$$

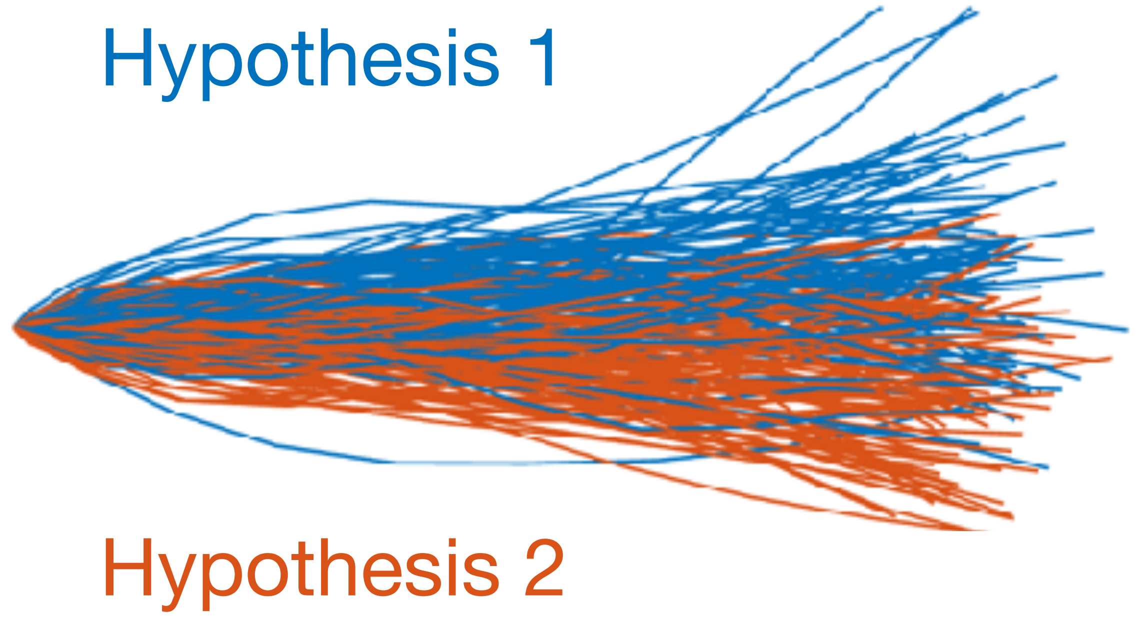
$$u^{F}_{t} \ \ x^{F}_{t} \sim \mathcal{N}(\mu_t, \Sigma_t)$$

What Makes Inference Easy/Difficult?

If $M = M^i$, dynamic programming shows:

$$x_t^{\mathrm{F}} \sim \mathcal{N}(\xi_t^i, \Lambda_t) \quad \begin{cases} q_t^i = (E_t^{\mathrm{F}})^\top q_{t+1}^i - Q^{\mathrm{F}} M^i x_t^{\mathrm{L}} \\ \xi_{t+1}^i = E_t^{\mathrm{F}} \xi_t^i - F_t^{\mathrm{F}} q_{t+1}^i \\ \Lambda_{t+1} = E_t^{\mathrm{F}} \Lambda_t (E_t^{\mathrm{F}})^\top + F_t^{\mathrm{F}} \end{cases}$$

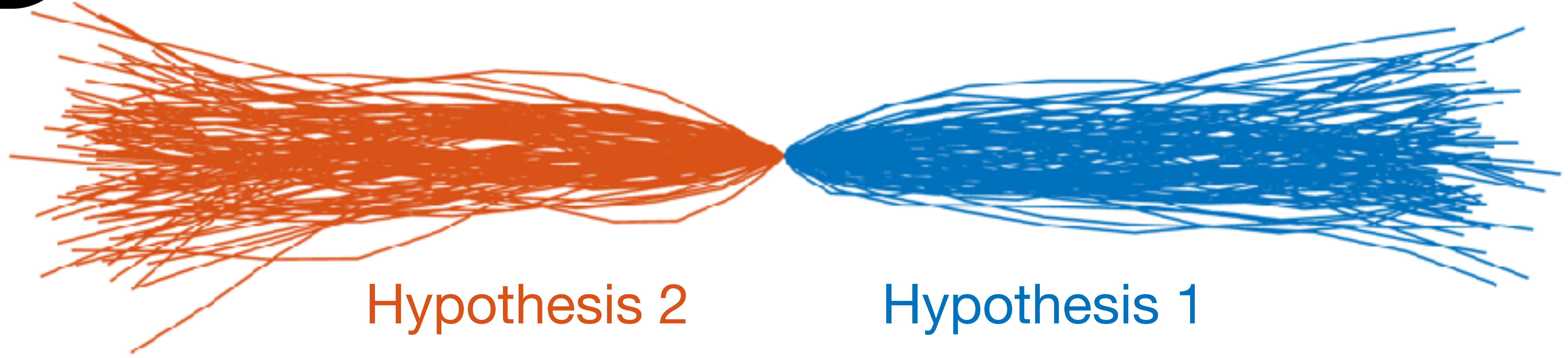
All depends on Leader!



Small Difference, Difficult Inference

KL-divergence between two distributions

$$D_{KL}\left(\xi_{0:\tau}^i, \xi_{0:\tau}^j, \Lambda_{0:\tau}\right) = \sum_{t=0}^{\tau} \left\| \xi_t^i - \xi_t^j \right\|_{\Lambda_t^{-1}}^2$$



Big Difference, Easy Inference

Maximizing Differences in Follower's Responses

Leader's input cost

$$\underset{u^{\mathtt{L}}_{0:\tau-1},\, x^{\mathtt{L}}_{0:\tau}}{\text{minimize}} \quad -\min_{i\neq j}\left\{\sum_{t=0}^{\tau} \left\| \xi_t^i - \xi_t^j \right\|^2_{(\Lambda_t)^{-1}}\right\} + g(u^{\mathtt{L}}_{0:\tau-1})$$

$$\text{subject to} \quad x^{\mathtt{L}}_{t+1} = A^{\mathtt{L}} x^{\mathtt{L}}_t + B^{\mathtt{L}} u^{\mathtt{L}}_t,\ u^{\mathtt{L}} \in \mathbb{U}$$

$$q_t^i = (E_t^{\mathtt{F}})^\top q_{t+1}^i - Q^{\mathtt{F}} M^i x_t^{\mathtt{L}}$$

$$\xi_{t+1}^i = E_t^{\mathtt{F}} \xi_t^i - F_t^{\mathtt{F}} q_{t+1}^i$$

$$i = 1,2,\ldots d$$

Leader's trajectory w/ constraints

Worst-case KL-divergence

(Difference of convex functions)

Follower's hypothetical trajectories

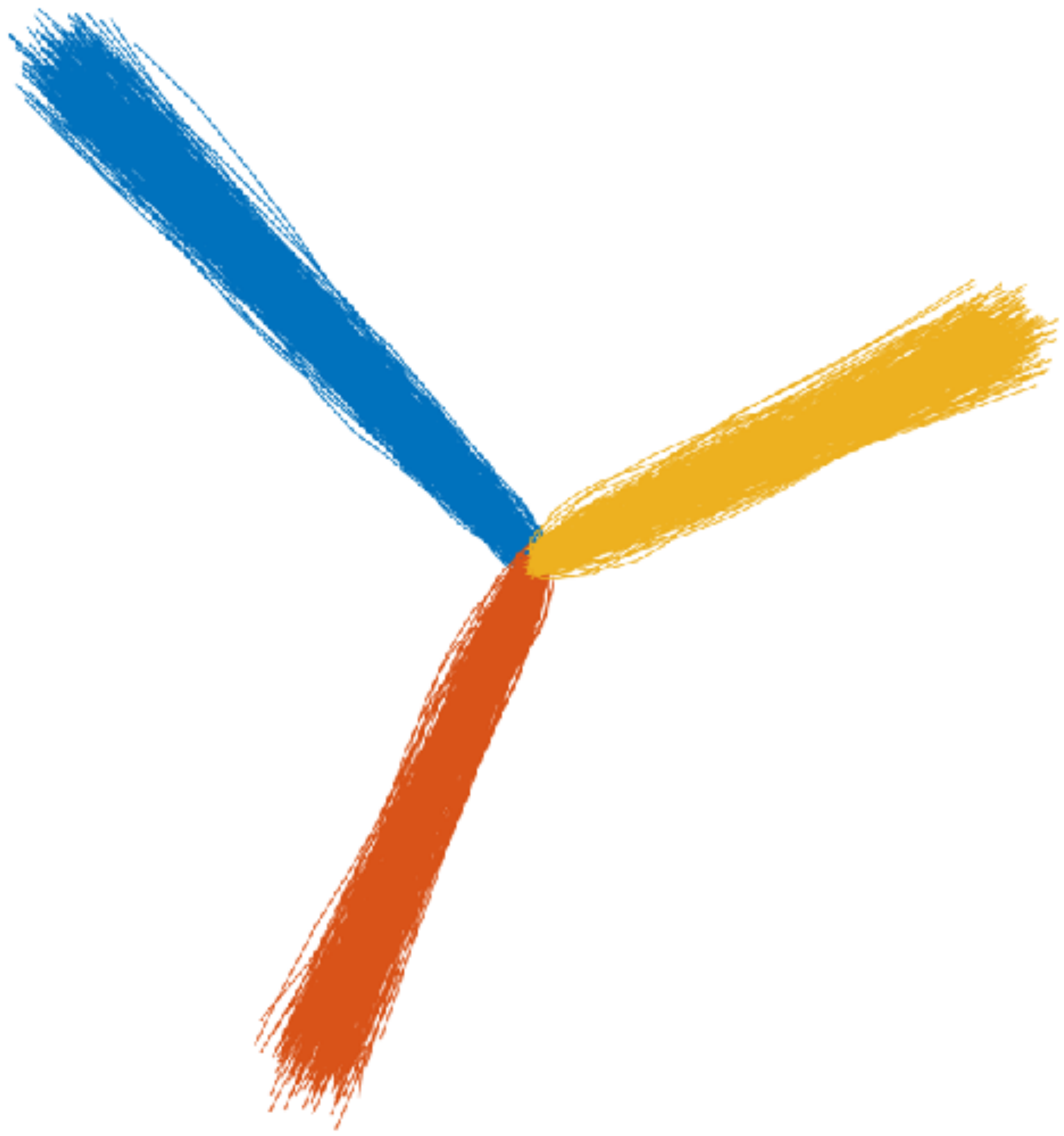
(linear functions of leader's trajectory)

We can solve it efficiently using linearization!

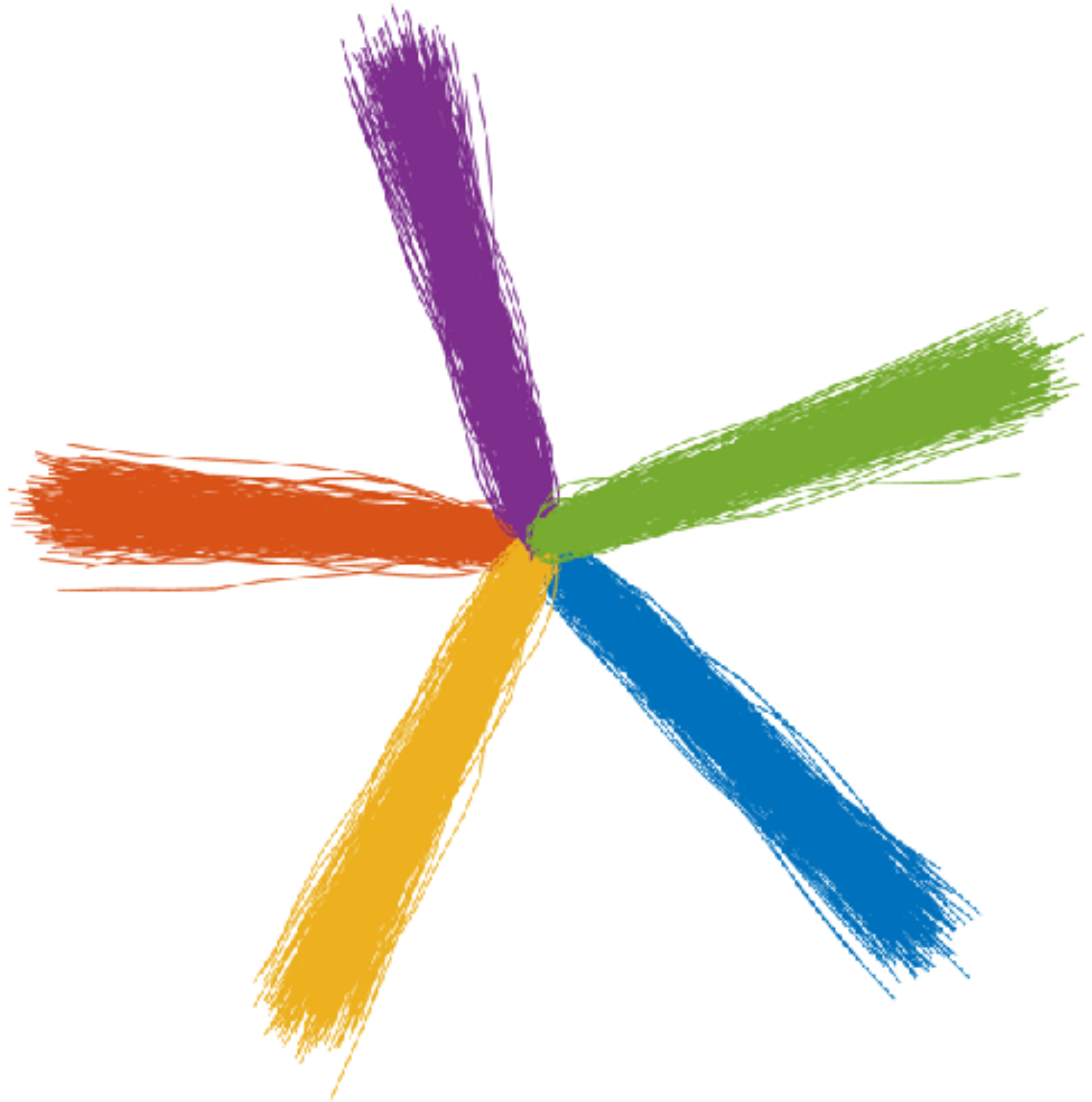
Numerical Experiments: Multiple Targets vs One Chaser

- Leader controls multiple agents, Follower controls one single agent
- Leader knows that Follower is chasing one agent, but not which one

3 leading agents



5 leading agents



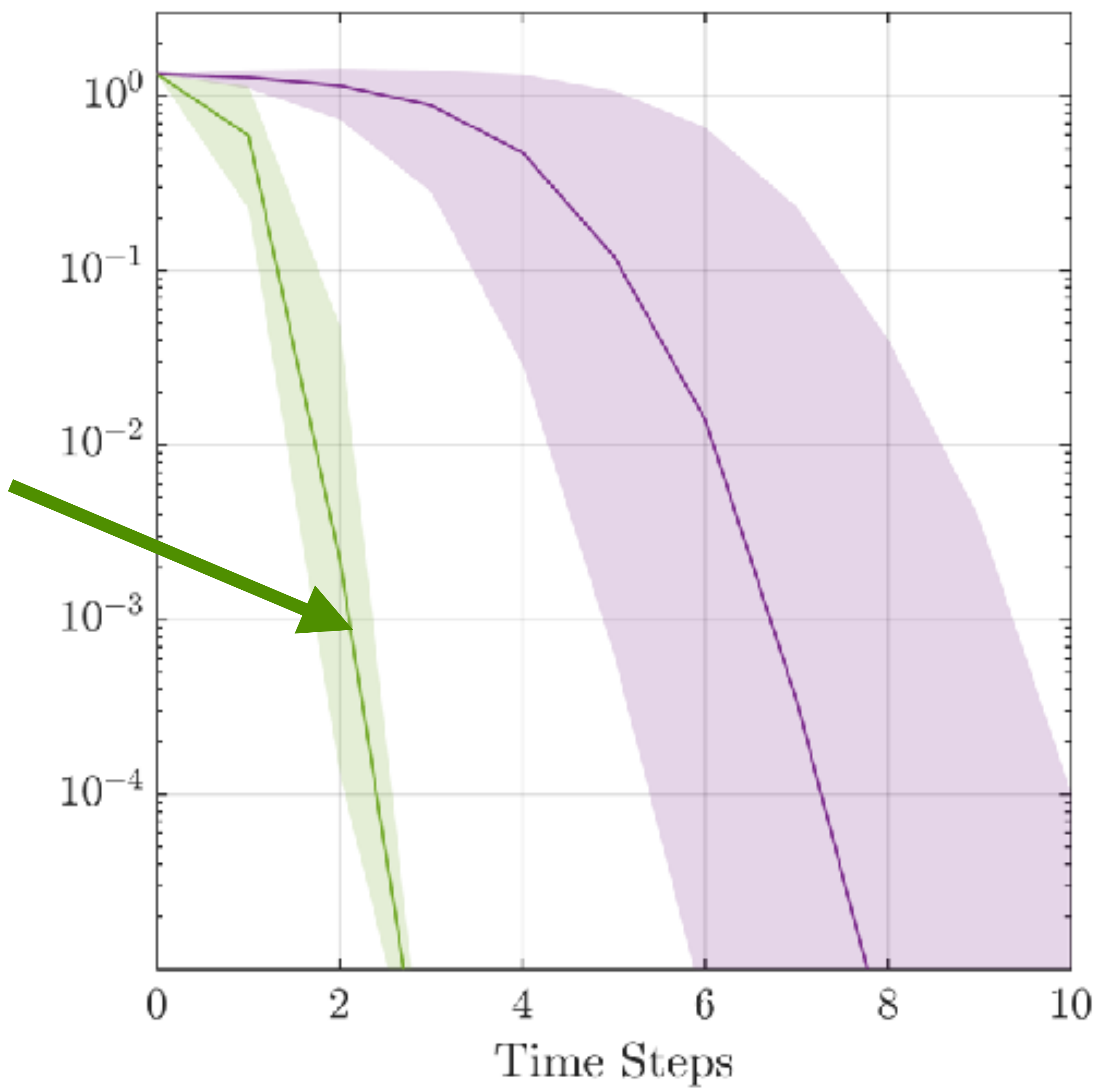
Follower's trajectory distributions under different hypothesis

Numerical Experiments: Multiple Targets vs One Chaser

- Leader controls multiple agents, Follower controls one single agent
- Leader knows that Follower is chasing one agent, but not which one

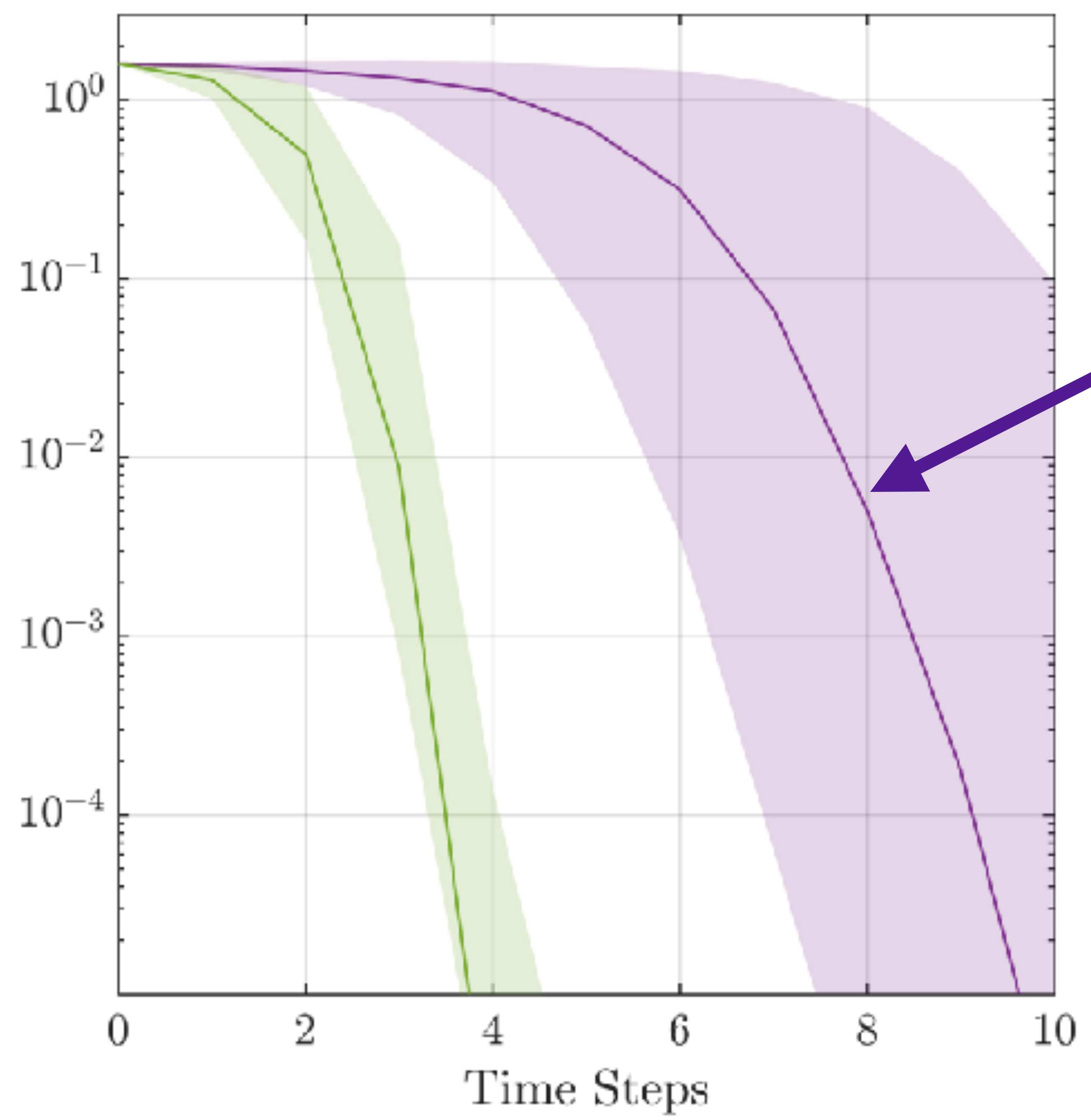
3 leading agents

proposed trajectories



5 leading agents

random trajectories



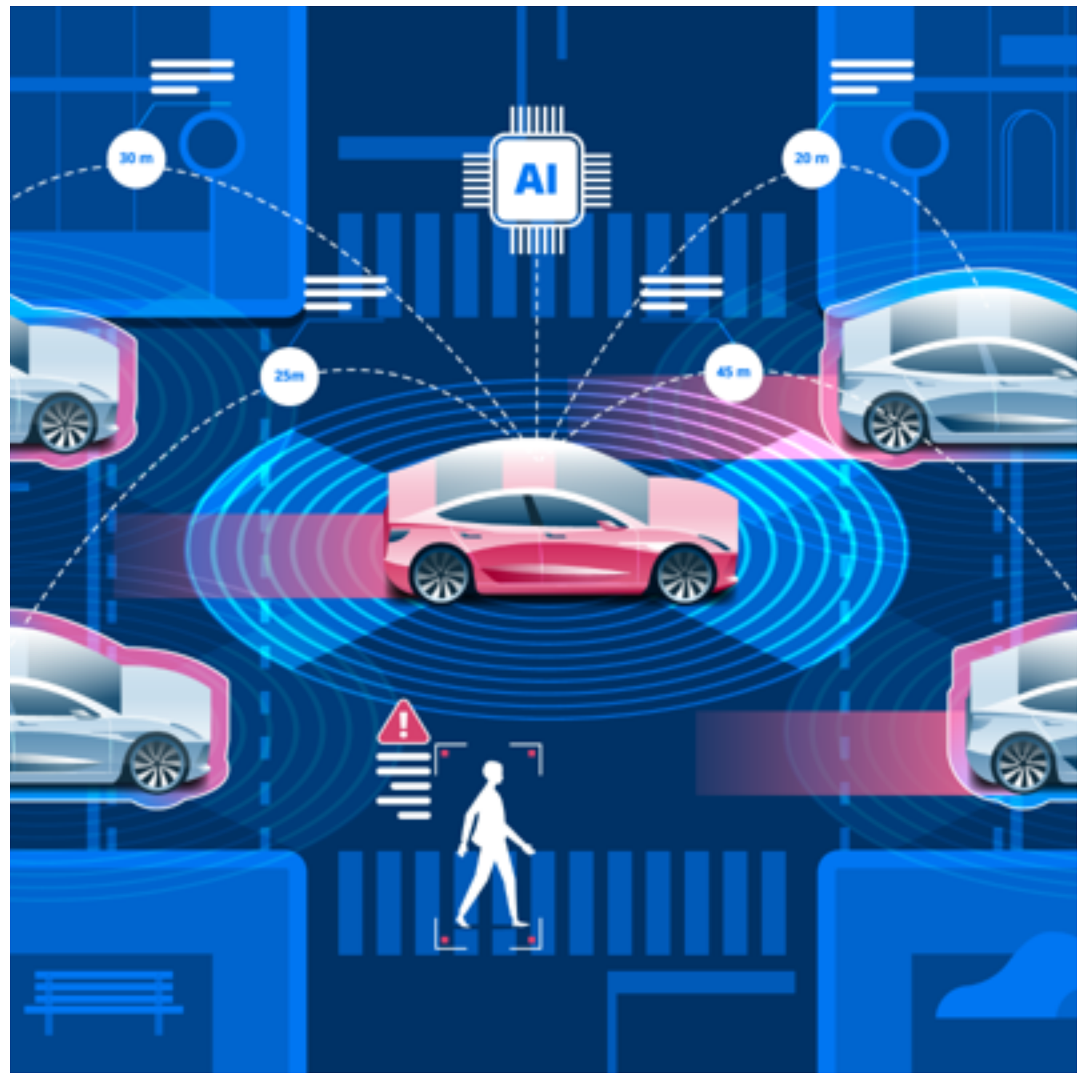
Total-variation error in leader's belief when using Bayesian learning

Future Directions

Shared Autonomy

futurebridge.com

Mixed Autonomy



fastdata.io

Cyberattacks & Defense

secplicity.org