Mixing Continuous Strategies: A Case Study in Trajectory Games

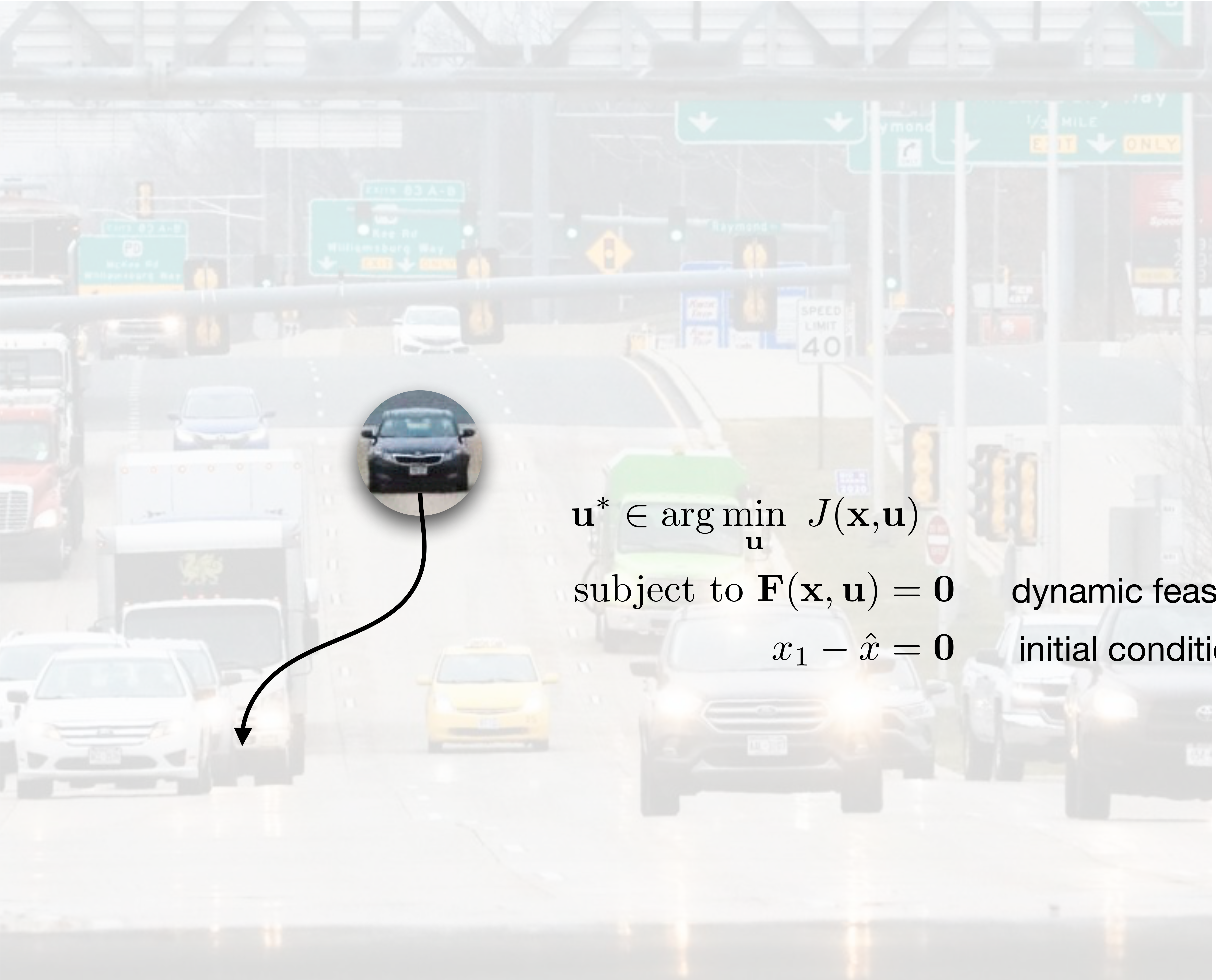
David Fridovich-Keil
Assistant Professor, UT Austin

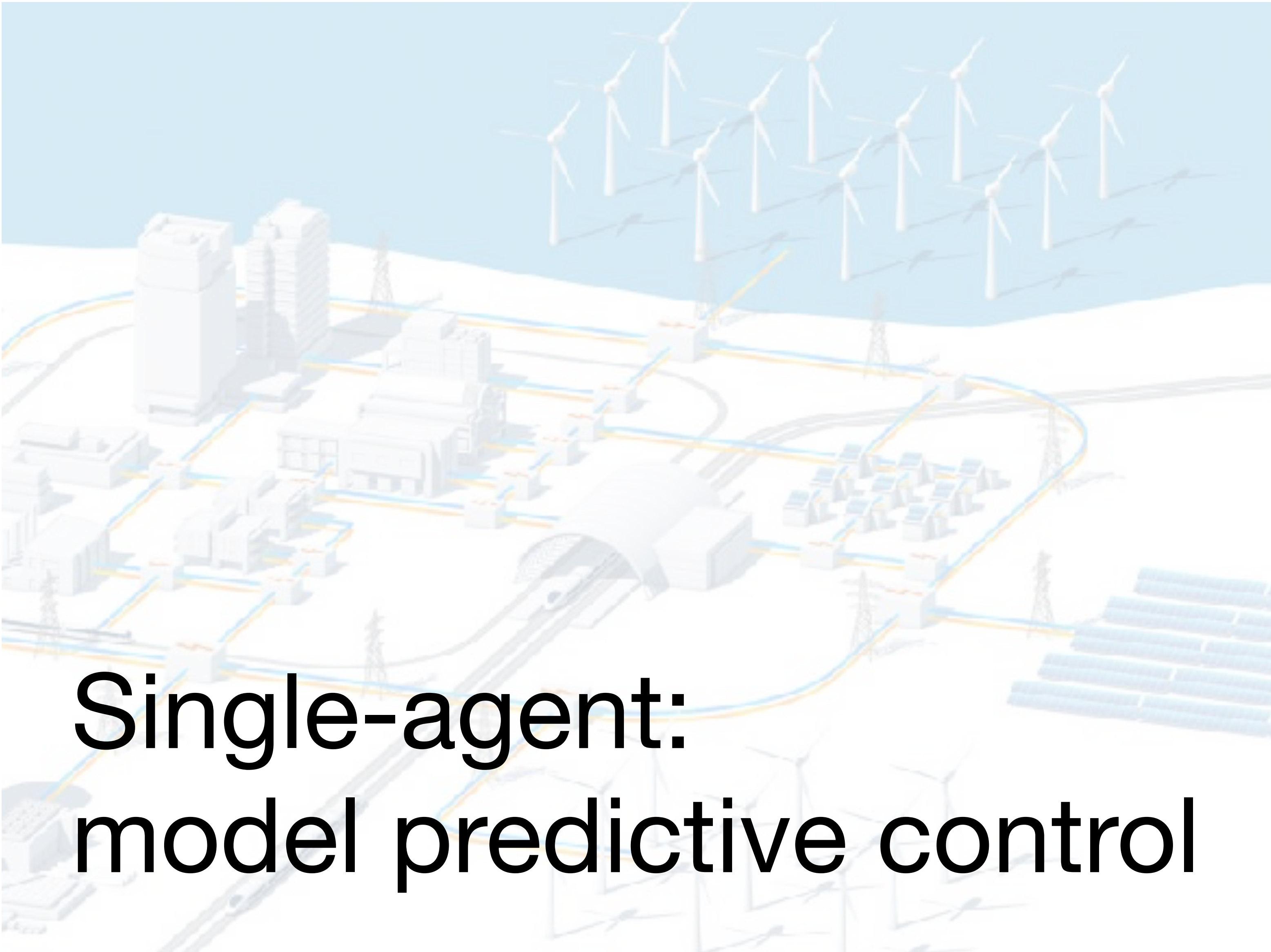
October 2022

Control and Learning for Autonomous Robotics


EXITS 83 A-B
McKee Rd
Williamsburg Way
EXIT ONLY
1/3 MILE
Raymond
SPEED LIMIT 40

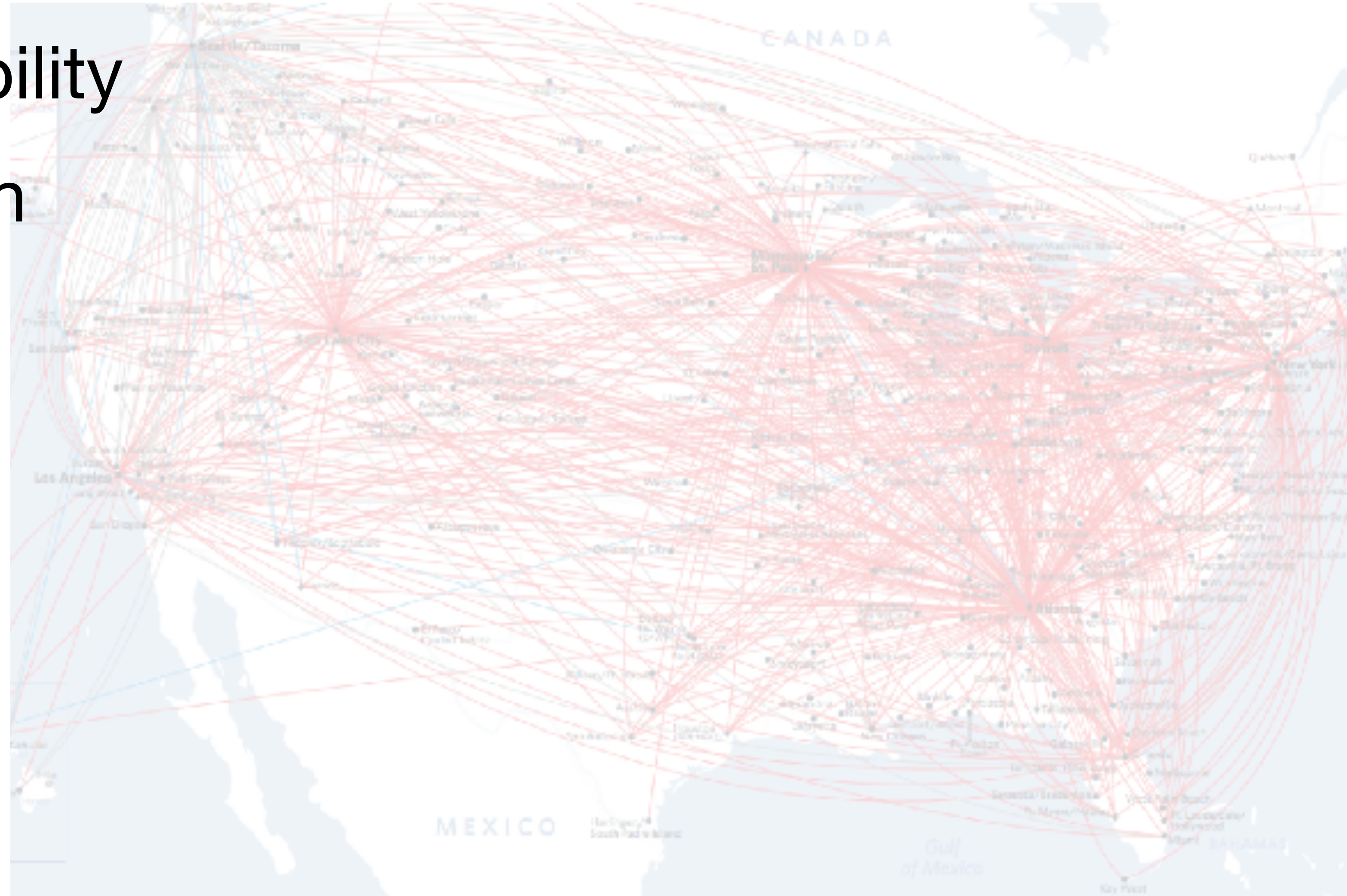

CANADA
MEXICO
Gulf of Mexico



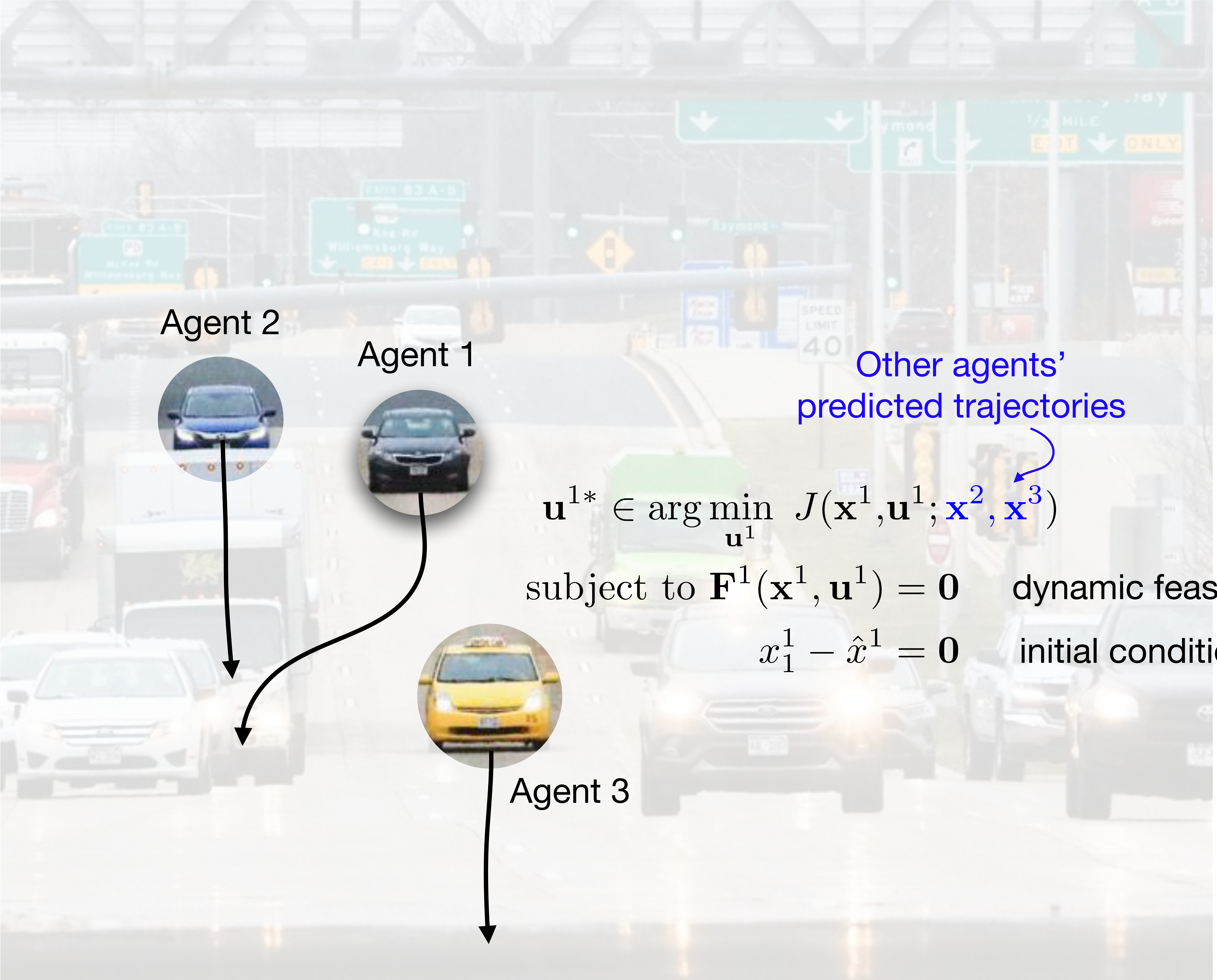


Single-agent: model predictive control

$$
\begin{aligned}
\mathbf{u}^* \in \arg\min_{\mathbf{u}} \;\; & J(\mathbf{x},\mathbf{u}) \\
\text{subject to } \; & \mathbf{F}(\mathbf{x},\mathbf{u}) = \mathbf{0} && \text{dynamic feasibility} \\
& x_1 - \hat{x} = \mathbf{0} && \text{initial condition}
\end{aligned}
$$



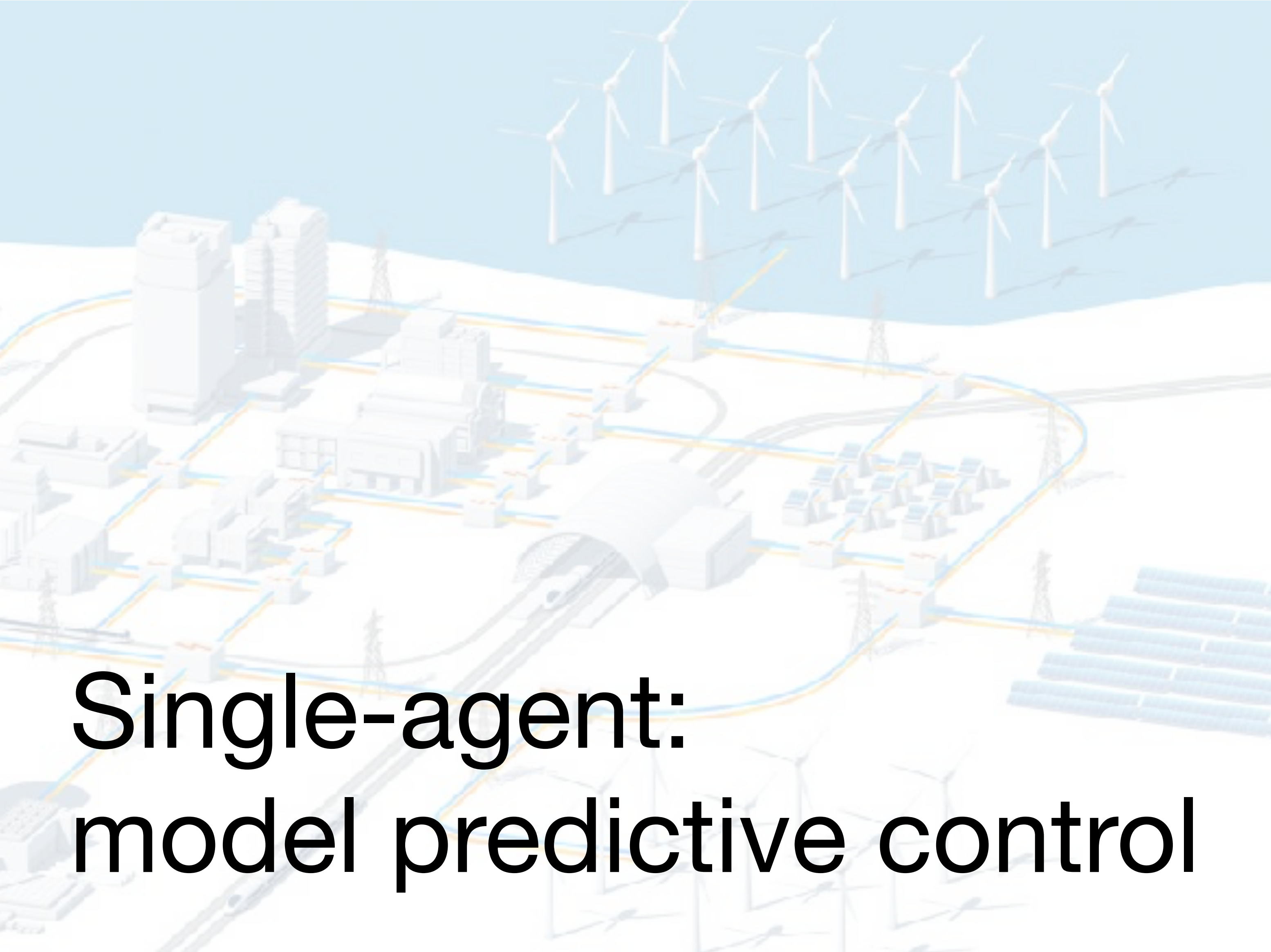
Single-agent: model predictive control

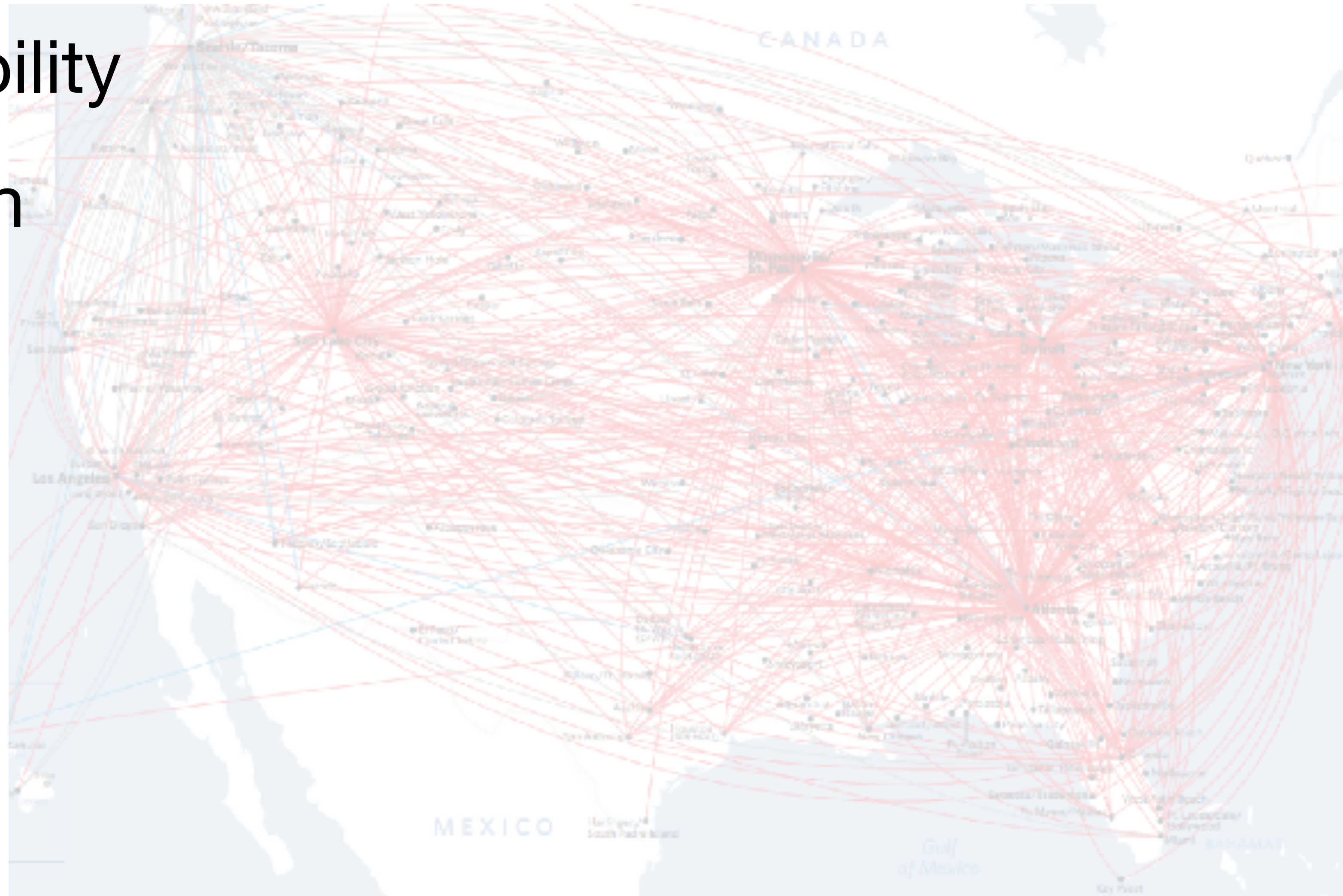


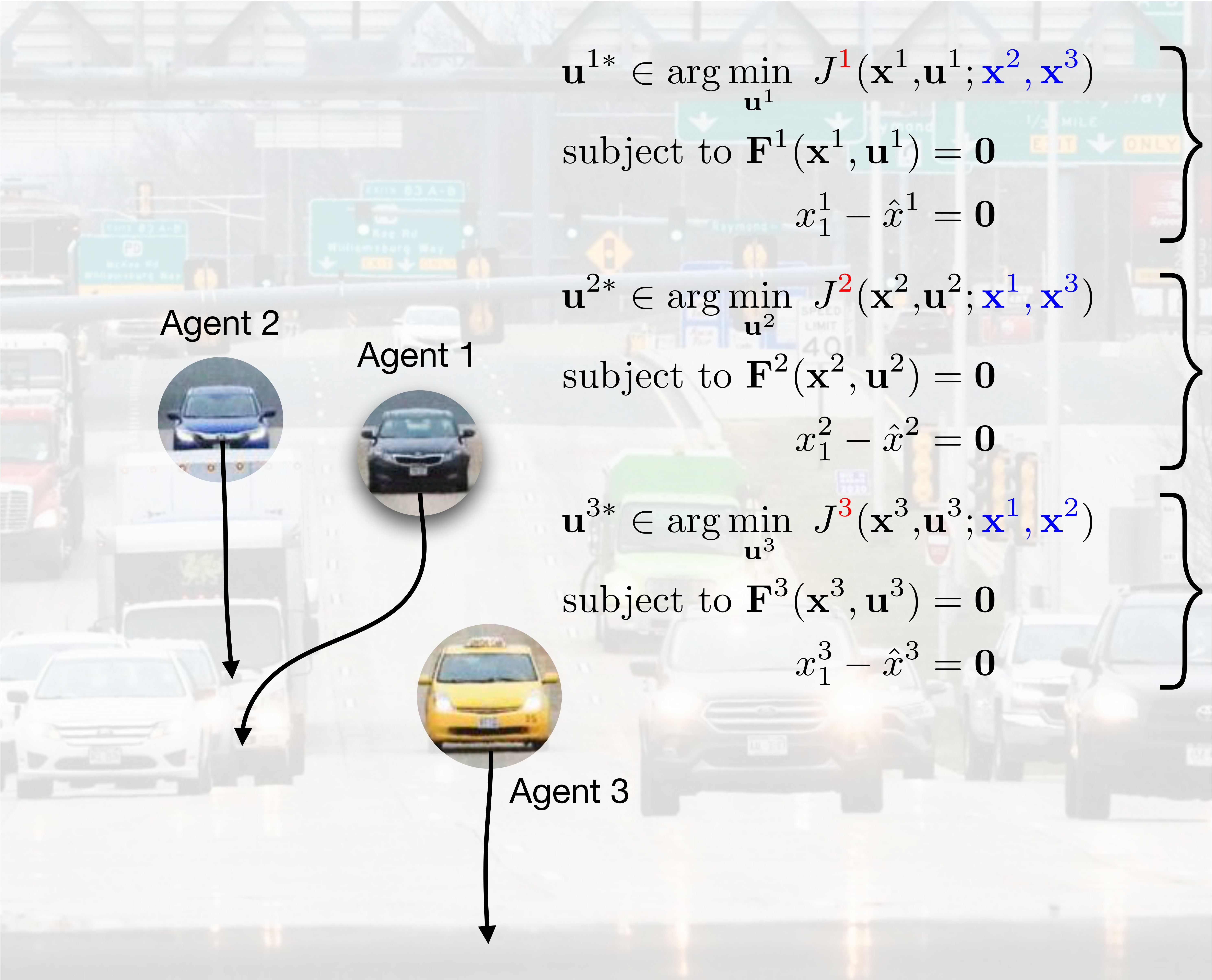
$$\mathbf{u}^{1*} \in \arg\min_{\mathbf{u}^1} \; J(\mathbf{x}^1, \mathbf{u}^1; \mathbf{x}^2, \mathbf{x}^3)$$

$$\text{subject to } \mathbf{F}^1(\mathbf{x}^1, \mathbf{u}^1) = \mathbf{0} \qquad \text{dynamic feasibility}$$

$$x_1^1 - \hat{x}^1 = \mathbf{0} \qquad \text{initial condition}$$







$$\mathbf{u}^{1*} \in \arg\min_{\mathbf{u}^1} \ J^{1}(\mathbf{x}^1, \mathbf{u}^1; \mathbf{x}^2, \mathbf{x}^3)$$
$$\text{subject to } \mathbf{F}^1(\mathbf{x}^1, \mathbf{u}^1) = \mathbf{0}$$
$$x_1^1 - \hat{x}^1 = \mathbf{0}$$

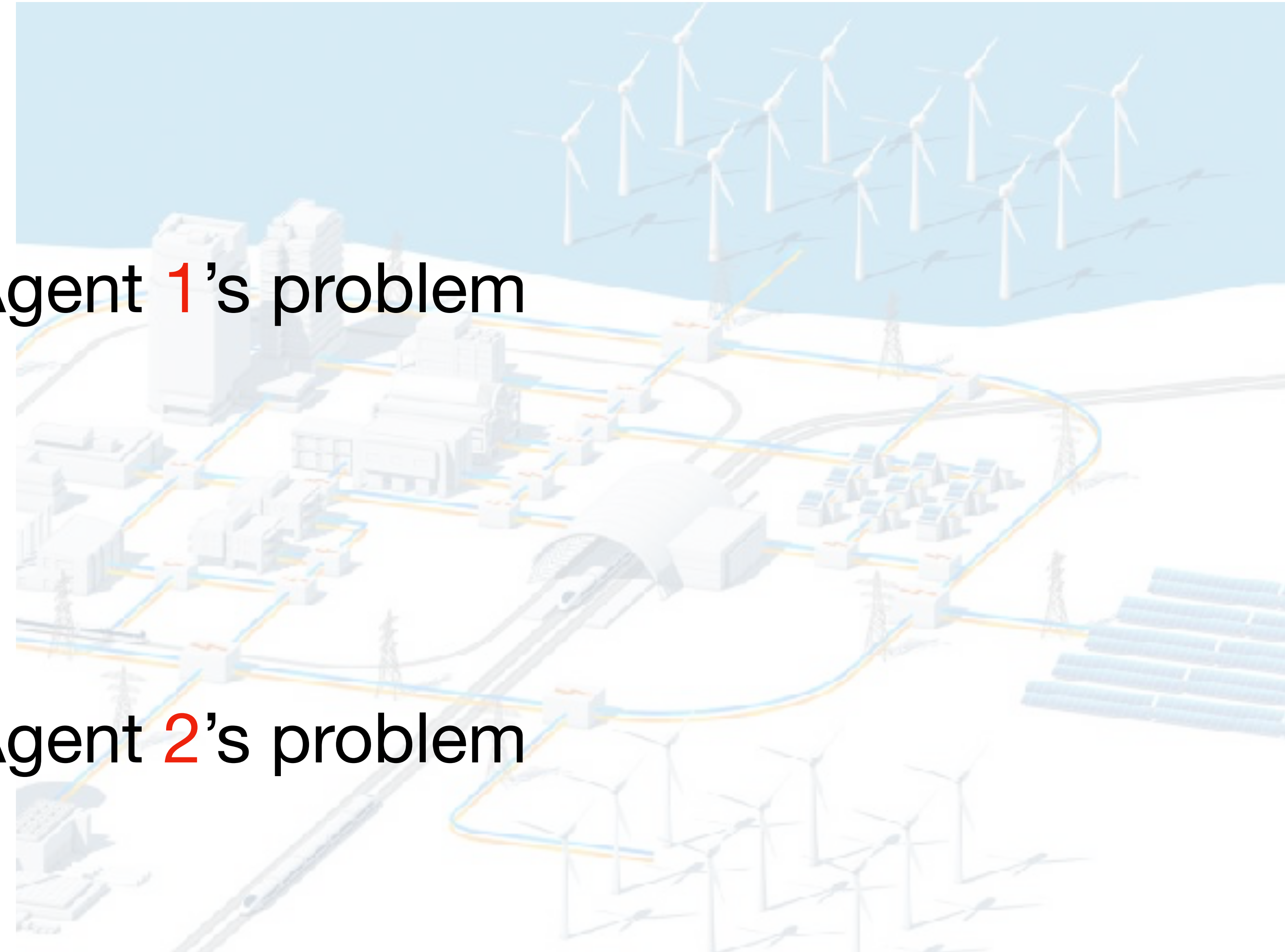
Agent 1's problem

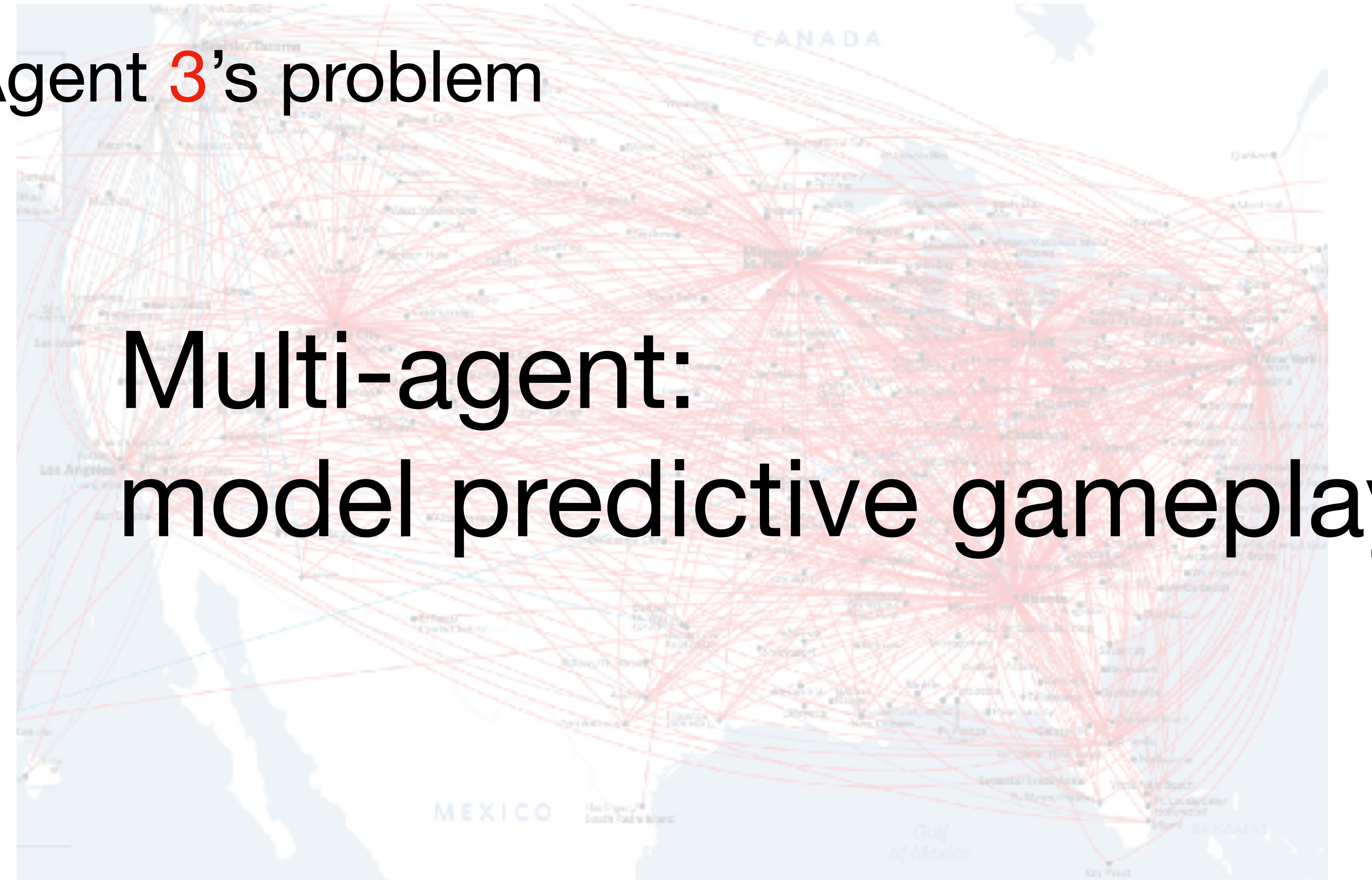
$$\mathbf{u}^{2*} \in \arg\min_{\mathbf{u}^2} \ J^{2}(\mathbf{x}^2, \mathbf{u}^2; \mathbf{x}^1, \mathbf{x}^3)$$
$$\text{subject to } \mathbf{F}^2(\mathbf{x}^2, \mathbf{u}^2) = \mathbf{0}$$
$$x_1^2 - \hat{x}^2 = \mathbf{0}$$

Agent 2's problem

$$\mathbf{u}^{3*} \in \arg\min_{\mathbf{u}^3} \ J^{3}(\mathbf{x}^3, \mathbf{u}^3; \mathbf{x}^1, \mathbf{x}^2)$$
$$\text{subject to } \mathbf{F}^3(\mathbf{x}^3, \mathbf{u}^3) = \mathbf{0}$$
$$x_1^3 - \hat{x}^3 = \mathbf{0}$$

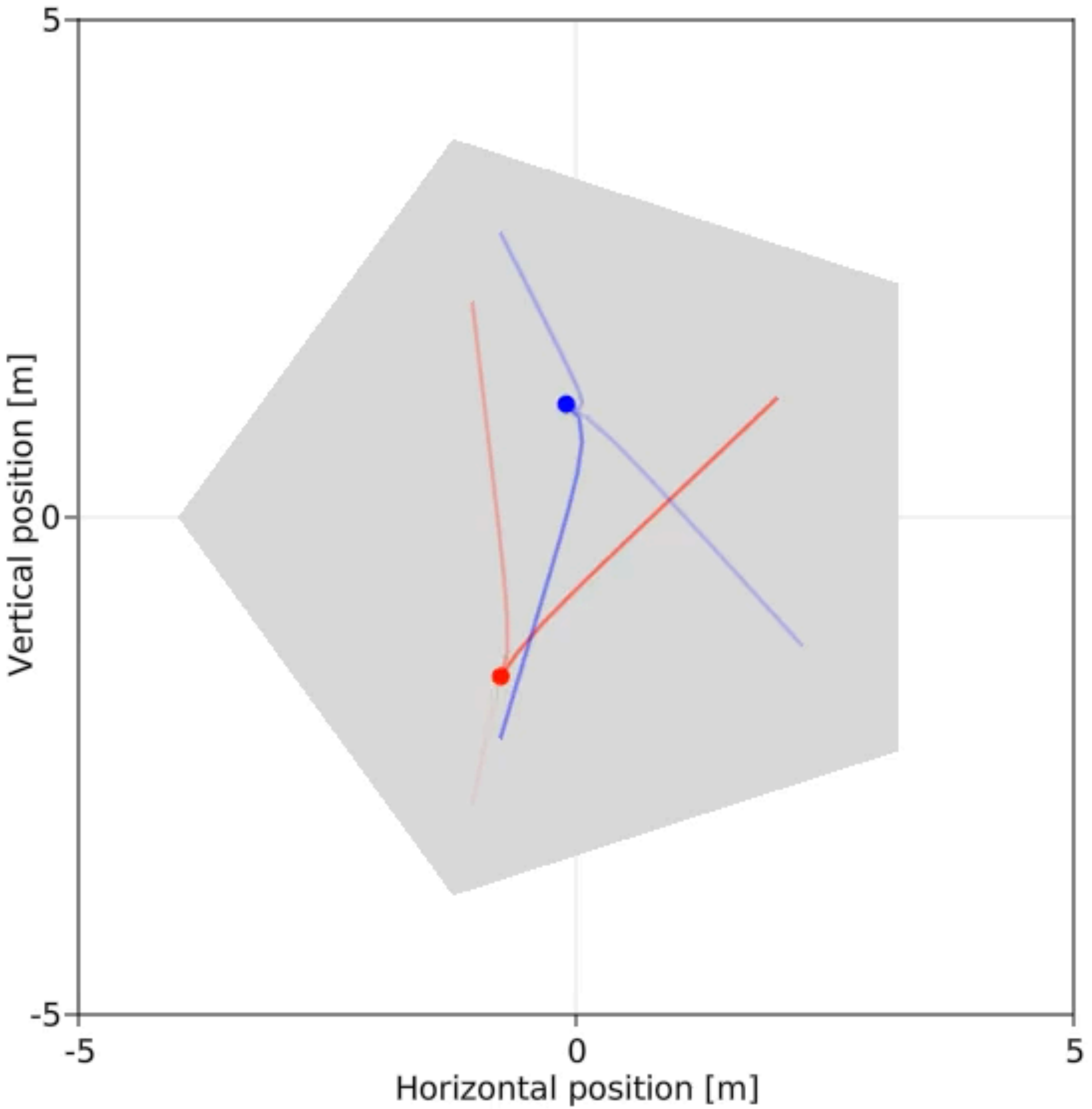
Agent 3's problem



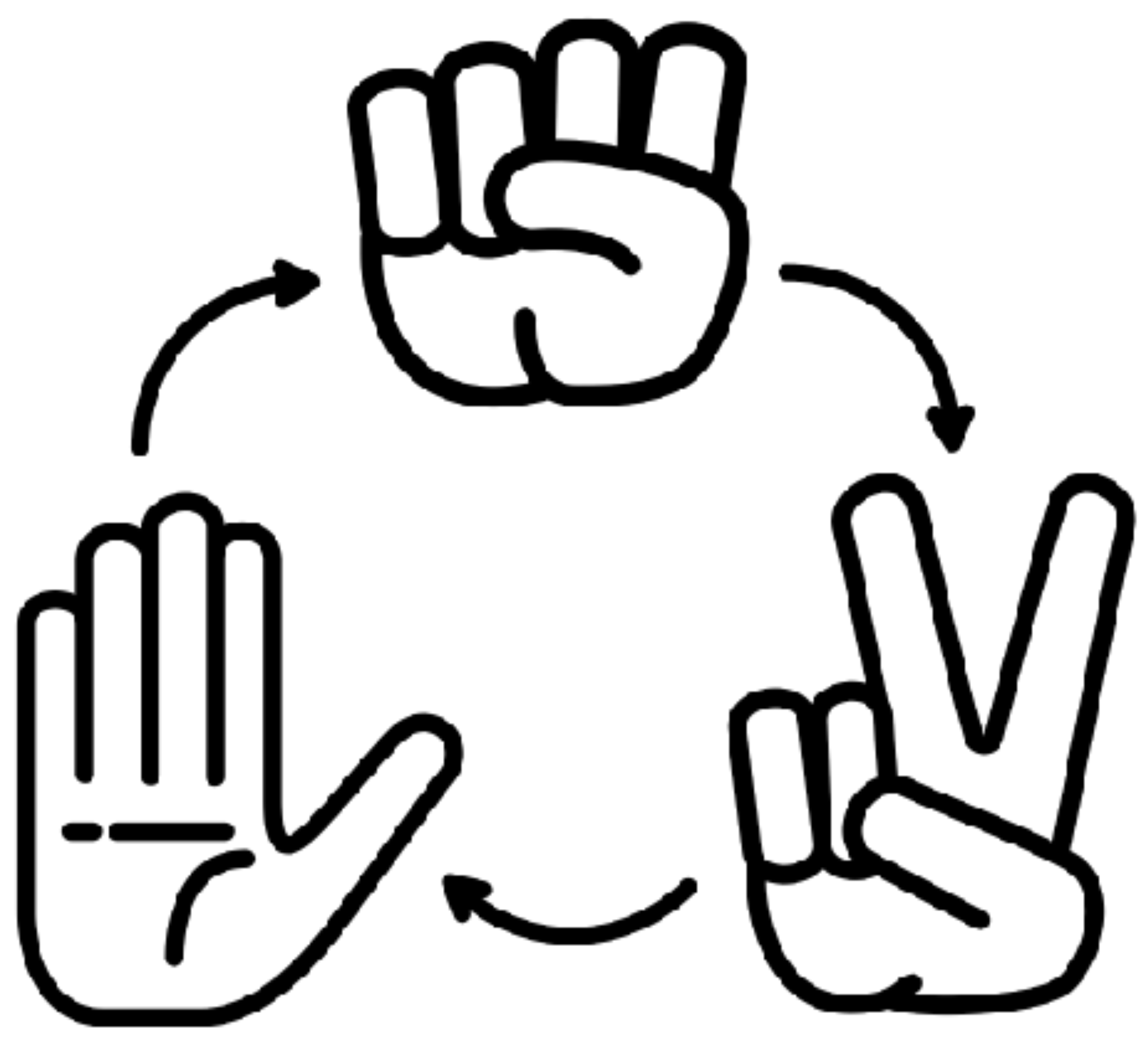


Multi-agent:
model predictive gameplay

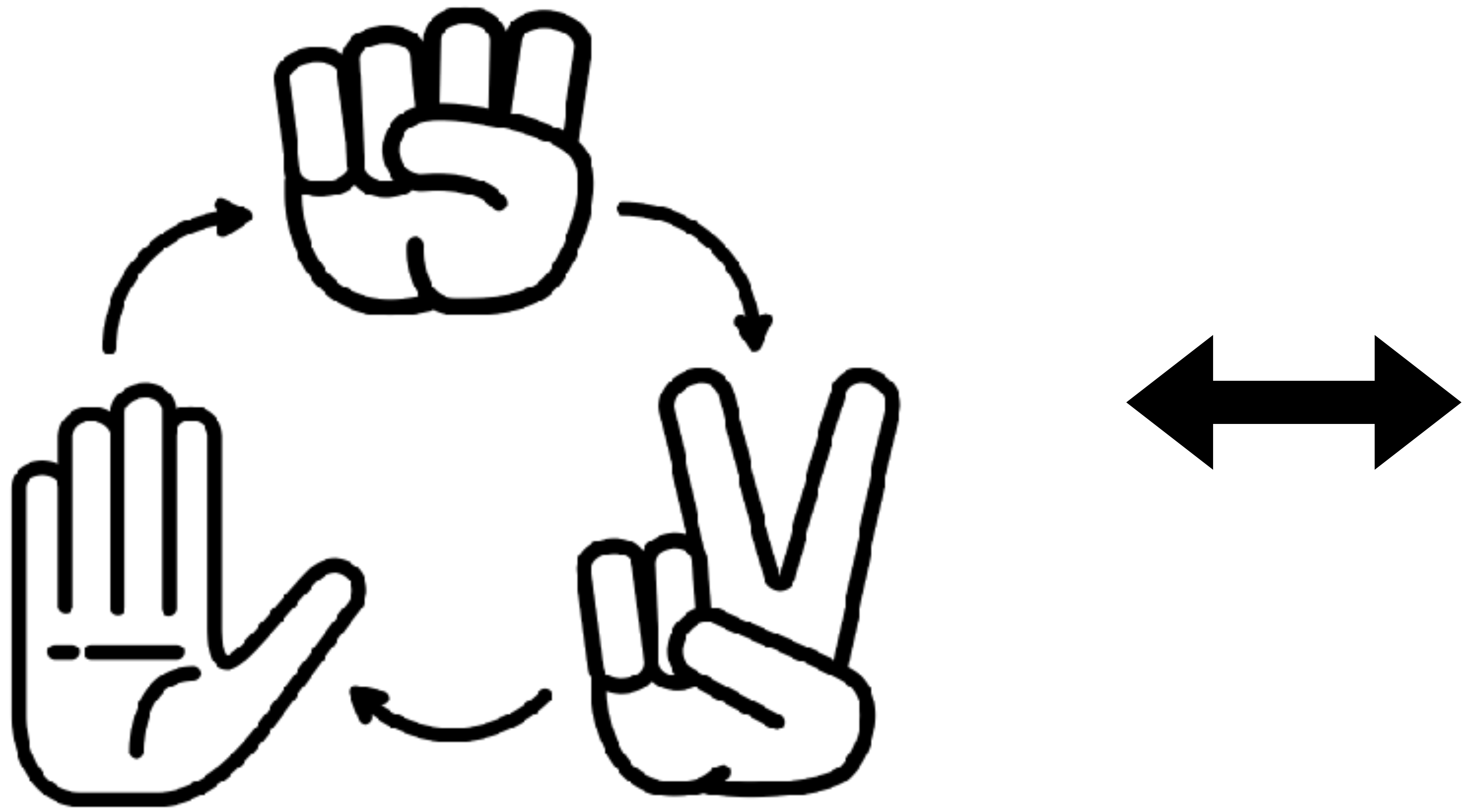
*Laine, **DFK**, et. al., "The Computation of Approximate Generalized Feedback Nash Equilibria" in SIOPT 2022.


5
0
-5
-5
0
5
Vertical position [m]
Horizontal position [m]

“Pure” Nash equilibria may not exist

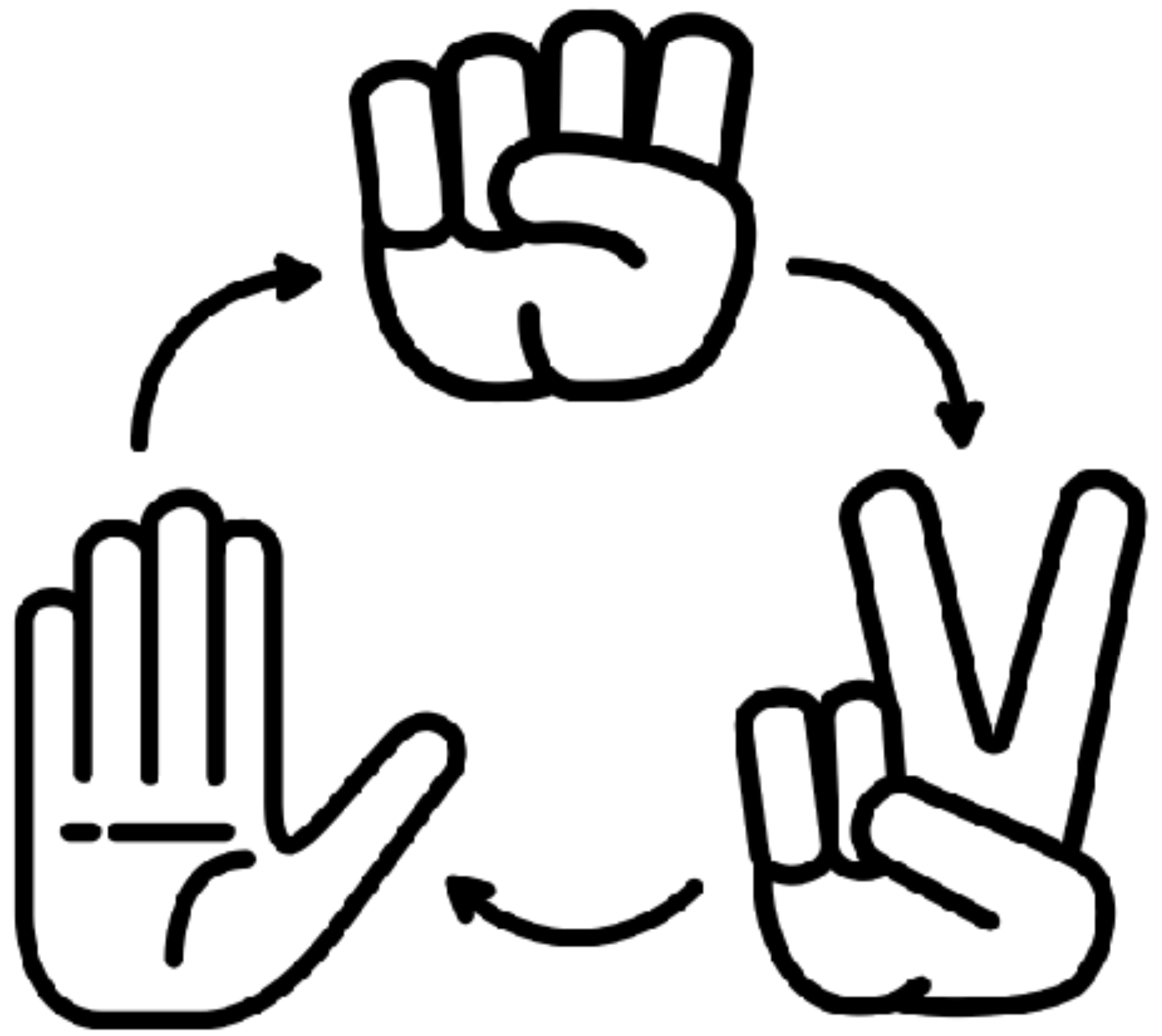


…but “mixed” Nash exist in finite spaces



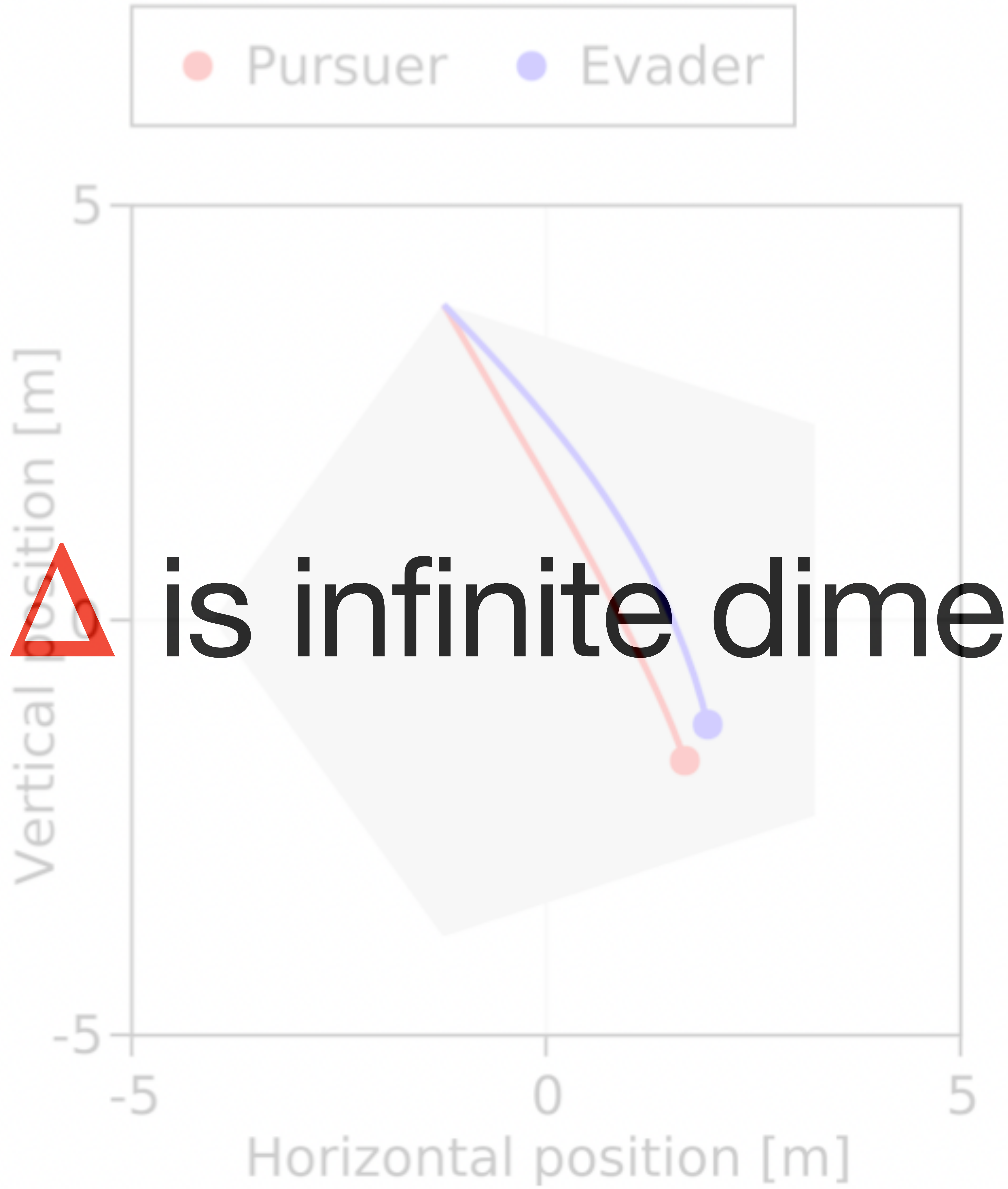
(Bi)matrix game $\begin{cases} \min_{p\in\Delta} \; p^\top A q \\ \min_{q\in\Delta} \; p^\top B q \end{cases}$

…but "mixed" Nash exist in finite spaces



(Bi)matrix game $\begin{cases} \min_{p \in \Delta} \; p^\top A q \\ \min_{q \in \Delta} \; p^\top B q \end{cases}$

What if Δ is infinite dimensional?

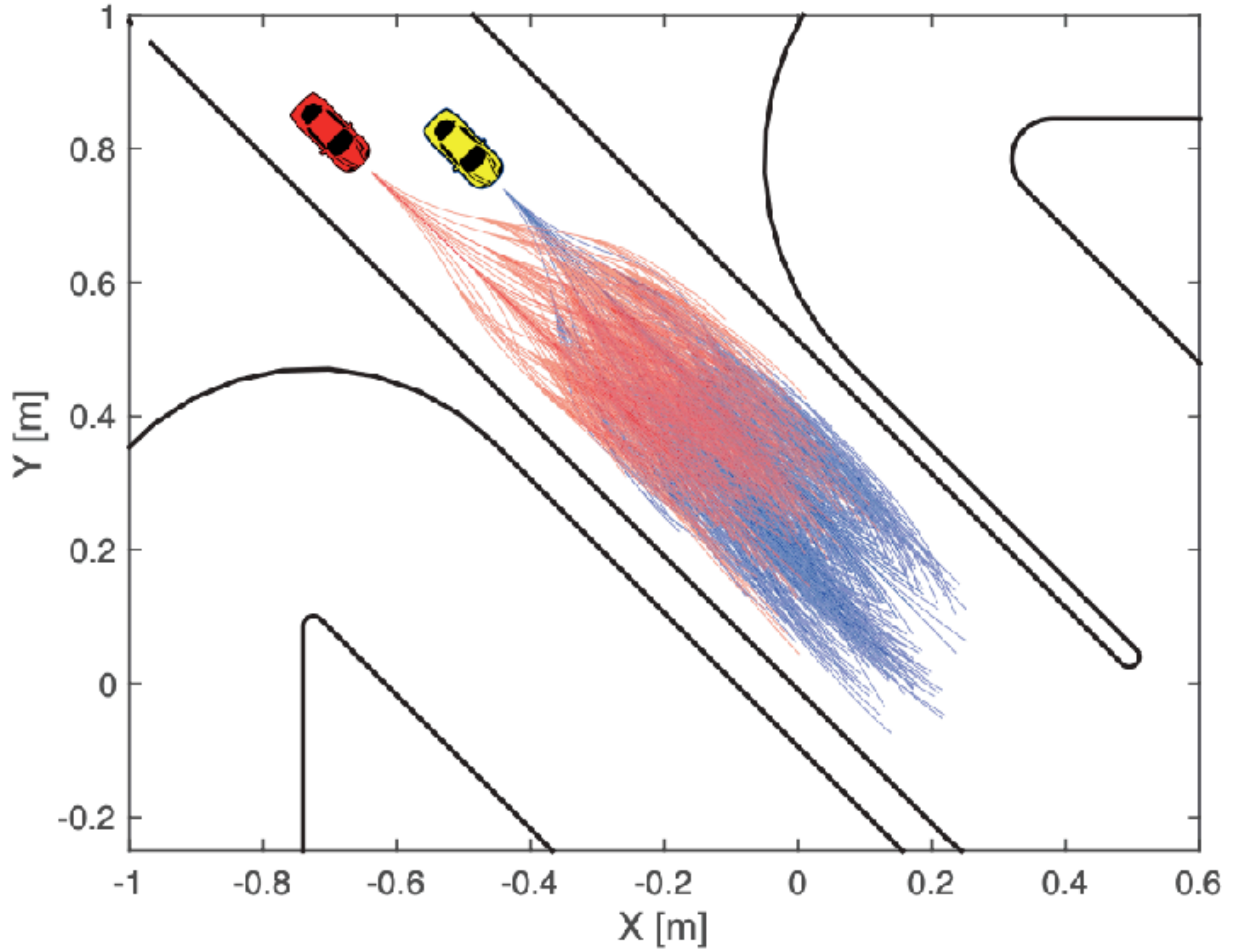


Approaching the infinite case…

- Can construct the τ_j^i by dense sampling*
  - Not super efficient… big matrix game

- Observation 1: Sampling more trajectories yields quantifiably better performance#

- Observation 2: Mixed Nash solutions are almost always sparse!



*A. Liniger and J. Lygeros, "A Noncooperative Game Approach to Autonomous Racing," in T-CST, 2020.

#Bopardikar et. al., "Randomized Sampling for Large Zero-Sum Games," in Automatica, 2013.

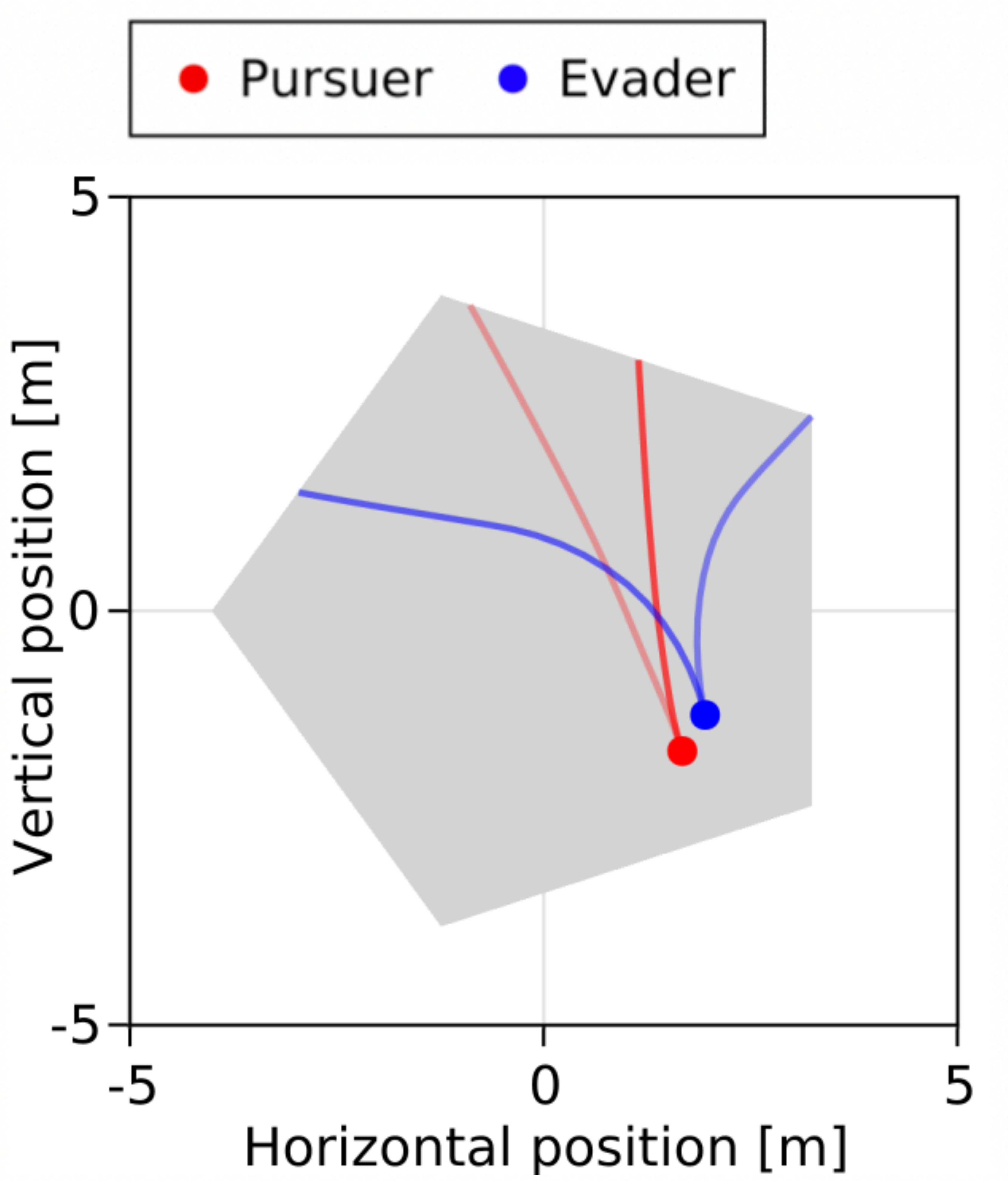
Idea: find sparse solutions directly via a two-stage game

A two-stage formulation*

$$\left.\begin{array}{ll}\min\limits_{\{\tau_1^1,\ldots,\tau_n^1\}\in\mathcal{T}_1^n} & p^{*\top} A q^* \\ \max\limits_{\{\tau_1^2,\ldots,\tau_m^2\}\in\mathcal{T}_2^m} & p^{*\top} A q^*\end{array}\right\}\text{ Upper stage}$$

$$\text{subject to } (p^*, q^*) \text{ a Nash equilibrium}$$

$$\text{where } A_{ij} = f(\tau_i^1, \tau_j^2)$$



*L. Peters, **DFK**, et. al., "Learning Mixed Strategies in Trajectory Games" in RSS 2022.

A two-stage formulation*

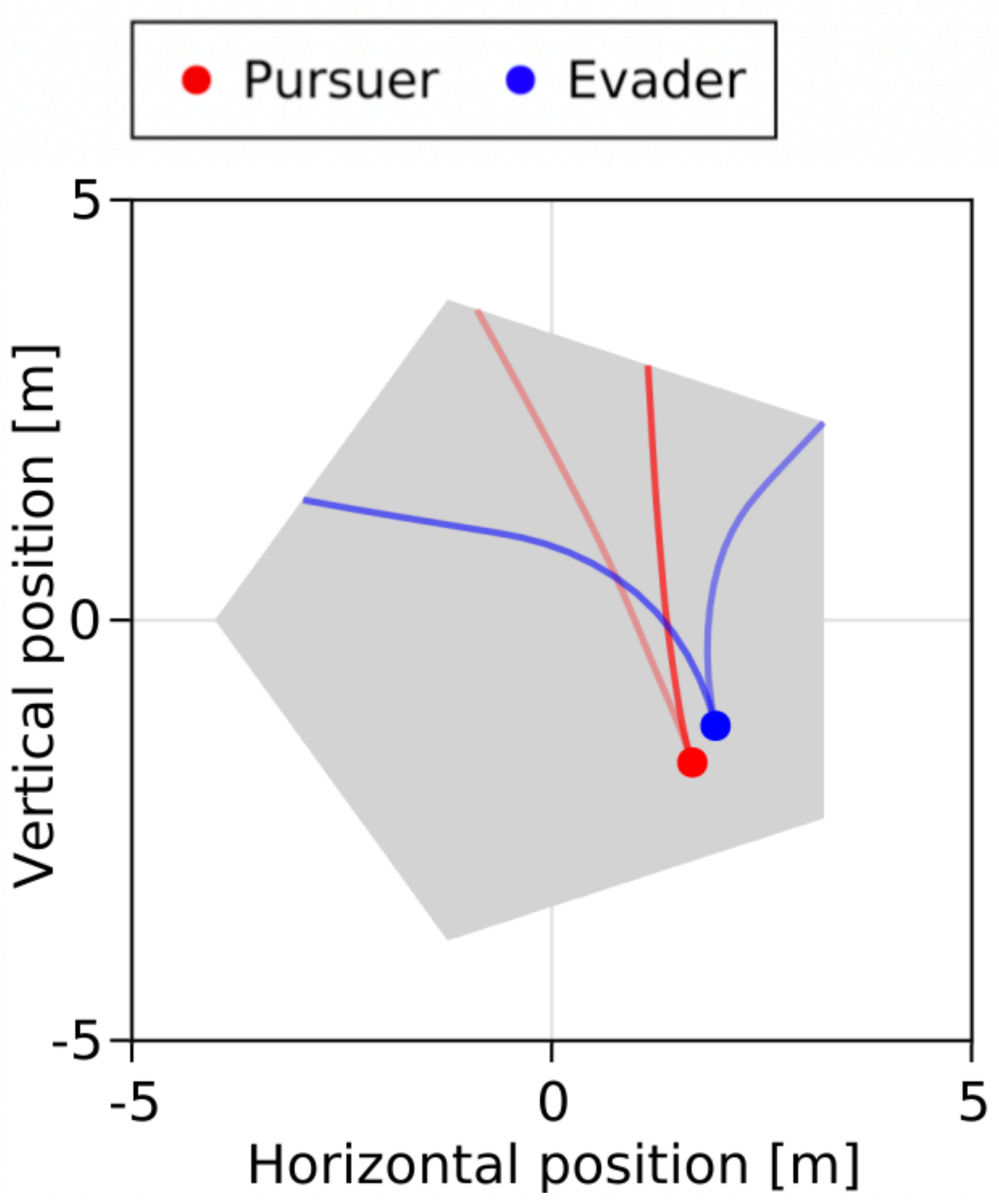
$$\min_{\{\tau_1^1,\ldots,\tau_n^1\}\in\mathcal{T}_1^n} \; p^{*\top} A q^* \quad \max_{\{\tau_1^2,\ldots,\tau_m^2\}\in\mathcal{T}_2^m} \; p^{*\top} A q^*$$

Upper stage

$$\text{subject to } \; p^* \in \arg\min_{p\in\Delta} \; p^\top A q \qquad q^* \in \arg\max_{q\in\Delta} \; p^\top A q$$

Lower stage

where $A_{ij} = f(\tau_i^1, \tau_j^2)$



*L. Peters, **DFK**, et. al., "Learning Mixed Strategies in Trajectory Games" in RSS 2022.

A gradient-based solution method

$$\begin{aligned} &\min_{\{\tau_1^1,\dots,\tau_n^1\}\in\mathcal{T}_1^n} \; p^{*\top} A q^* \\ &\max_{\{\tau_1^2,\dots,\tau_m^2\}\in\mathcal{T}_2^m} \; p^{*\top} A q^* \end{aligned} \quad \Bigg\} \text{ Upper stage}$$

$$\begin{aligned} \text{subject to } \; & p^* \in \arg\min_{p\in\Delta} \; p^\top A q \\ & q^* \in \arg\max_{q\in\Delta} \; p^\top A q \end{aligned} \quad \Bigg\} \text{ Lower stage}$$

Simultaneous gradient descent/ascent on $\tau^1_{1:n}$ and $\tau^2_{1:m}$…

…while chaining through the solution map at the lower level

L. Peters, **DFK, et. al., “Learning Mixed Strategies in Trajectory Games” in RSS 2022.*

Implicit differentiation of NE

- Need to differentiate backward through (bi)matrix game solve
- Bimatrix game ~ linear complementarity program
- Can still differentiate solution map implicitly

$$(q_1, q_2) \text{ a NE for fixed } (A, B) \iff \begin{aligned} q_1 \geq \mathbf{0} \perp Aq_2 \geq \mathbf{1} \\ q_2 \geq \mathbf{0} \perp B^\top q_1 \geq \mathbf{1} \end{aligned}$$

$$\text{strict complimentarity} \implies \begin{aligned} \bar{A}\bar{q}_2 &= \mathbf{1} \\ \bar{B}^\top \bar{q}_1 &= \mathbf{1} \\ \hat{q}_2 &= \mathbf{0} \\ \hat{q}_1 &= \mathbf{0} \end{aligned}$$

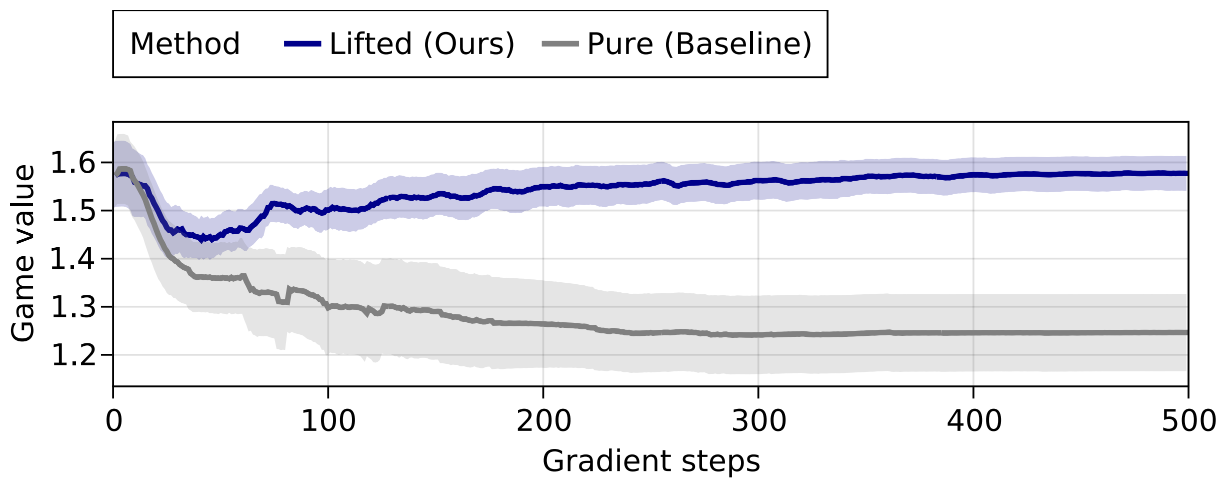
$$\frac{\partial \bar{q}_2}{\partial \bar{A}} = \frac{\partial}{\partial \bar{A}} \left[\bar{A}^{-1}\mathbf{1}\right]$$

only active constraints

Be careful if only weak complementarity holds!

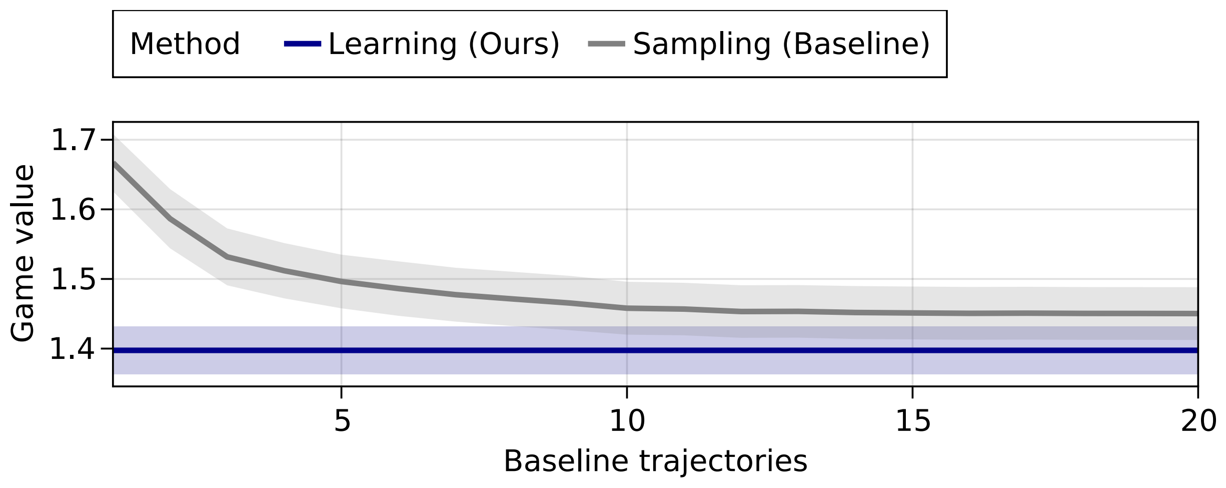
What happens?

Equilibrium characteristics



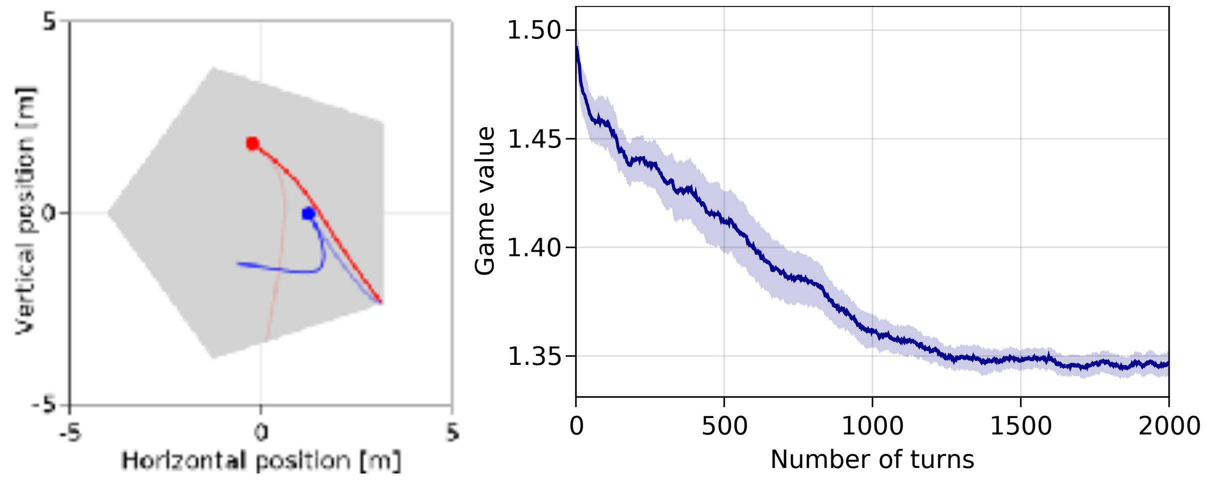
Lifting reliably converges to solutions which are NOT pure NE

Representative power



It's better to learn a few trajectory candidates than to sample many at random.

Online optimization in self-play



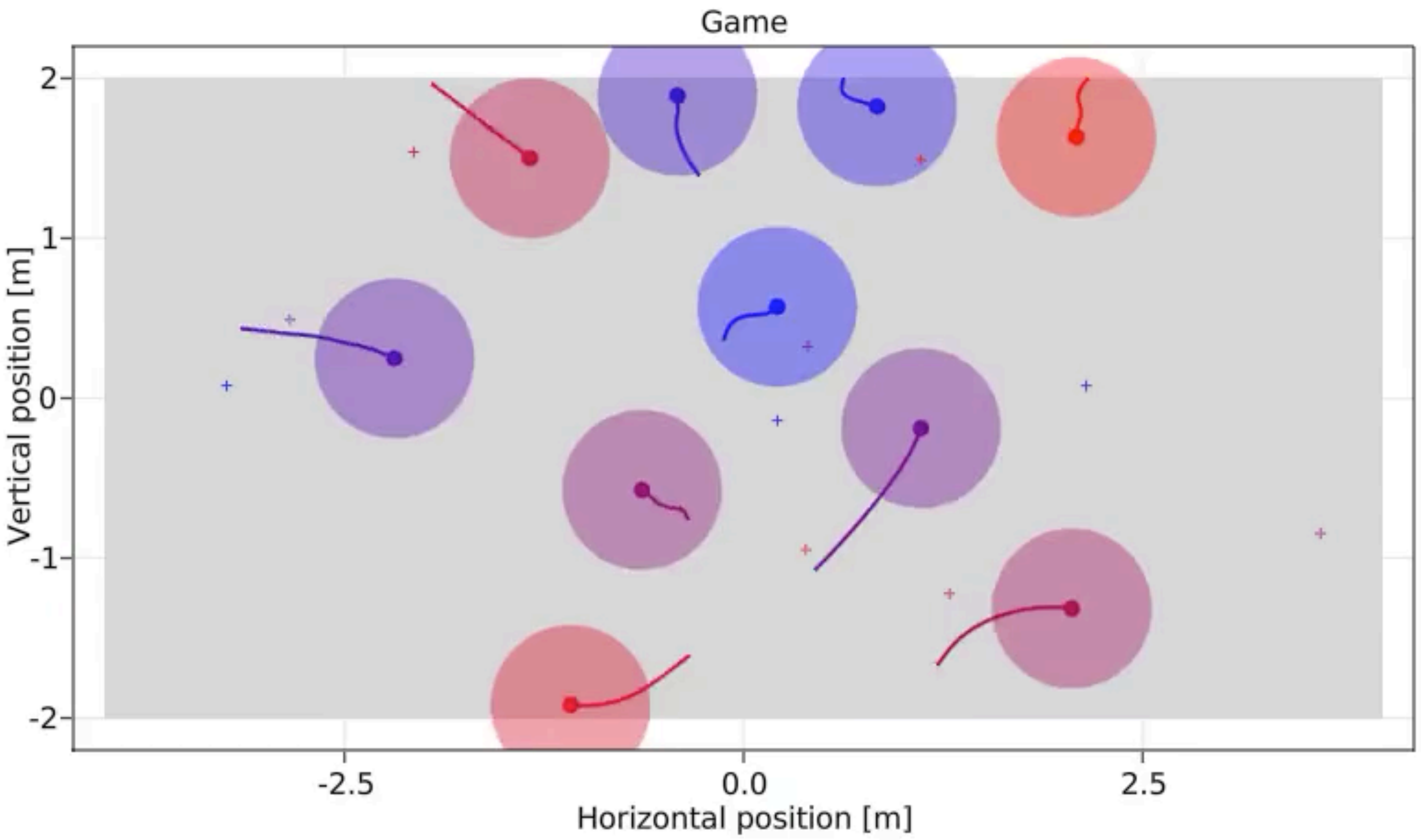
Lifted solvers can be pre-trained efficiently in self-play.

Meta-game: lifted vs. pure Nash

	Evader	
Pursuer	Lifted	Pure
Lifted	1.360 ± 0.003	1.289 ± 0.005
Pure	1.463 ± 0.004	0.903 ± 0.009

Lifted strategies reliably outperform pure strategies.

Epilogue: the N-player case



- N-player setting is not a bimatrix game (see TensorGames.jl)
  - 80ms to solve 6-player game with 3 actions each
- Parallelization + mixed-mode AD… pipeline scales very well for many players (single action):
  - 175ms per gradient step

Work with Maximilian Schmidt

Pure Nash solutions do not always exist in continuous trajectory games.
Mixed strategies depend upon fixed primitive trajectories.

Lifted strategies can be found efficiently via implicit differentiation,
and outperform pure counterparts.

Lasse Peters
TU Delft

Forrest Laine
Vanderbilt

Cyrill Stachniss
Bonn

Javier Alonso-Mora
TU Delft

Laura Ferranti
TU Delft

Thank you! Questions?