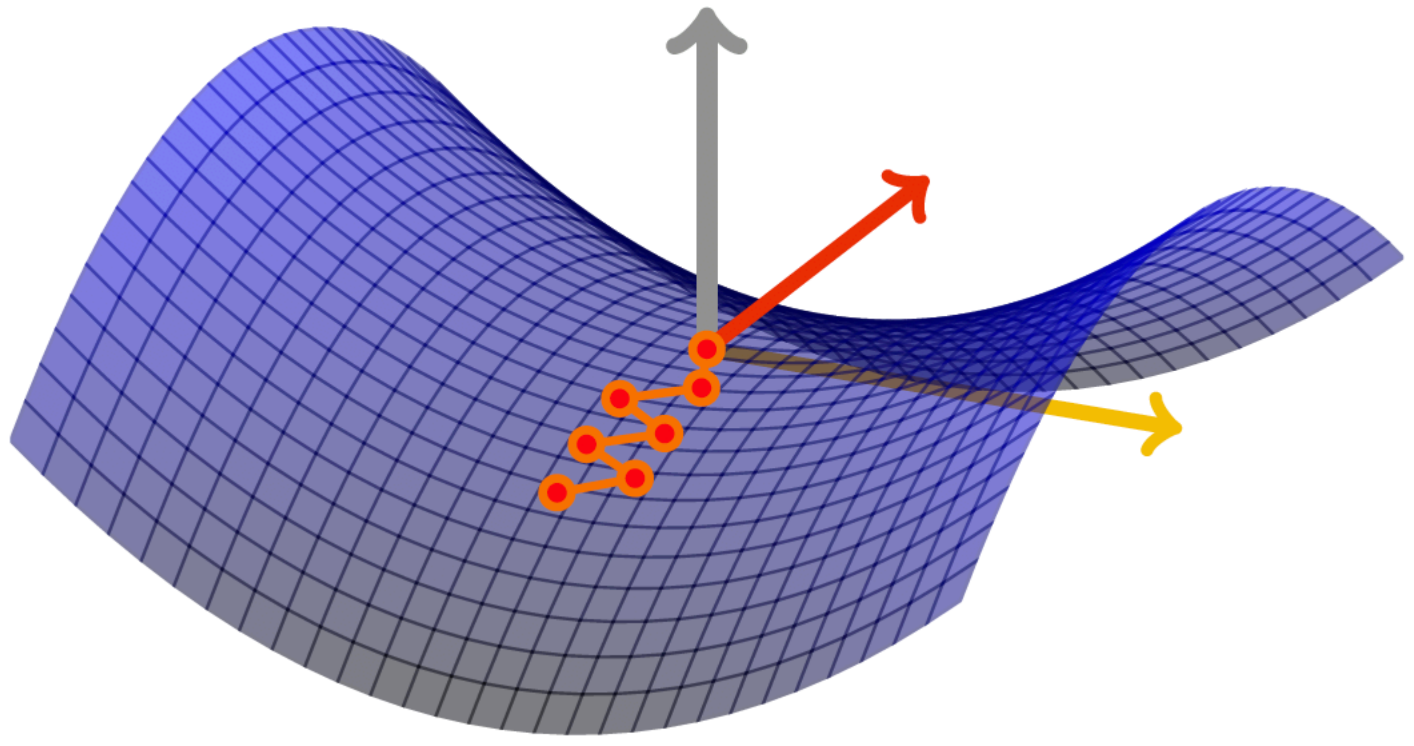
# Finite-time convergence of gradient-play in continuous games

**Benjamin Chasnov**, Lillian J. Ratliff, Daniel Calderone, Eric Mazumdar, and Samuel A. Burden

UNIVERSITY *of* WASHINGTON



# Can we solve **games** using tools from **optimization?**

# Can we solve **games** using tools from **optimization?**

Consider general-sum 2-player game:

$$x_1^\star = \arg\min_{x_1} \; f_1(x_1, x_2^\star)$$

$$x_2^\star = \arg\min_{x_2} \; f_2(x_1^\star, x_2)$$

# Can we solve **games** using tools from **optimization?**

Consider general-sum 2-player game:

$$x_1^\star = \arg\min_{x_1} \; f_1(x_1, x_2^\star)$$

$$x_2^\star = \arg\min_{x_2} \; f_2(x_1^\star, x_2)$$

Optimization of smooth $f(x)$ :

$$x^\star = \arg\min_{x} \; f(x)$$

# Can we solve **games** using tools from **optimization?**

Consider general-sum 2-player game:

$$x_1^\star = \arg\min_{x_1} \; f_1(x_1, x_2^\star)$$

$$x_2^\star = \arg\min_{x_2} \; f_2(x_1^\star, x_2)$$

Optimization of smooth $f(x)$:

$$x^\star = \arg\min_{x} \; f(x)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

# Can we solve **games** using tools from **optimization?**

Consider general-sum 2-player game:

$$x_1^\star = \arg\min_{x_1} \; f_1(x_1, x_2^\star)$$

$$x_2^\star = \arg\min_{x_2} \; f_2(x_1^\star, x_2)$$

Gradient-play, myopic updates:

$$x_1^+ \leftarrow x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ \leftarrow x_2 - \gamma D_2 f_2(x_1, x_2)$$

$\longleftrightarrow$ **?**

Optimization of smooth $f(x)$:

$$x^\star = \arg\min_{x} \; f(x)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

# Can we solve **games** using tools from **optimization?**

Consider general-sum 2-player game:

$$x_1^\star = \arg\min_{x_1} \; f_1(x_1, x_2^\star)$$

$$x_2^\star = \arg\min_{x_2} \; f_2(x_1^\star, x_2)$$

Gradient-play, myopic updates:

$$x_1^+ \leftarrow x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ \leftarrow x_2 - \gamma D_2 f_2(x_1, x_2)$$

?

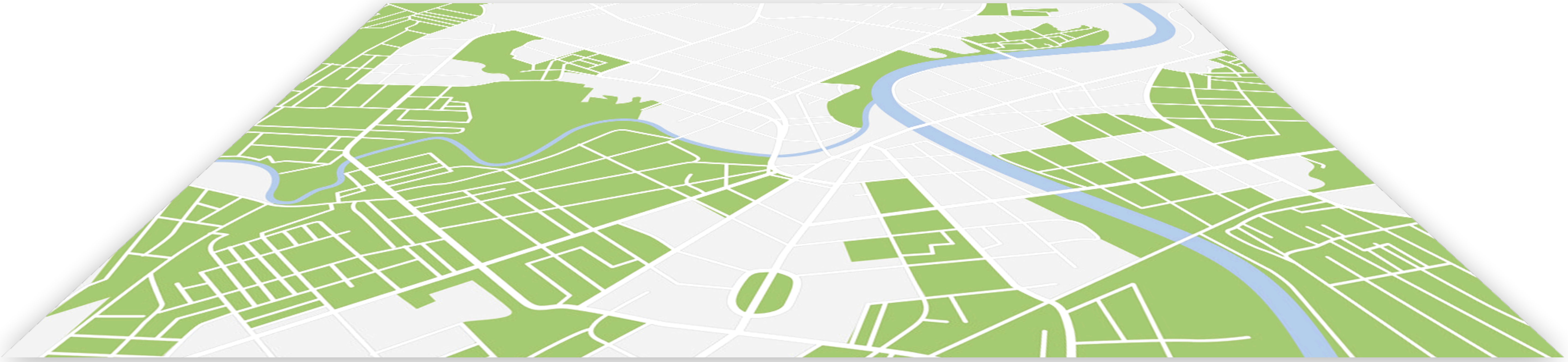
Optimization of smooth $f(x)$:

$$x^\star = \arg\min_{x} \; f(x)$$

Gradient descent

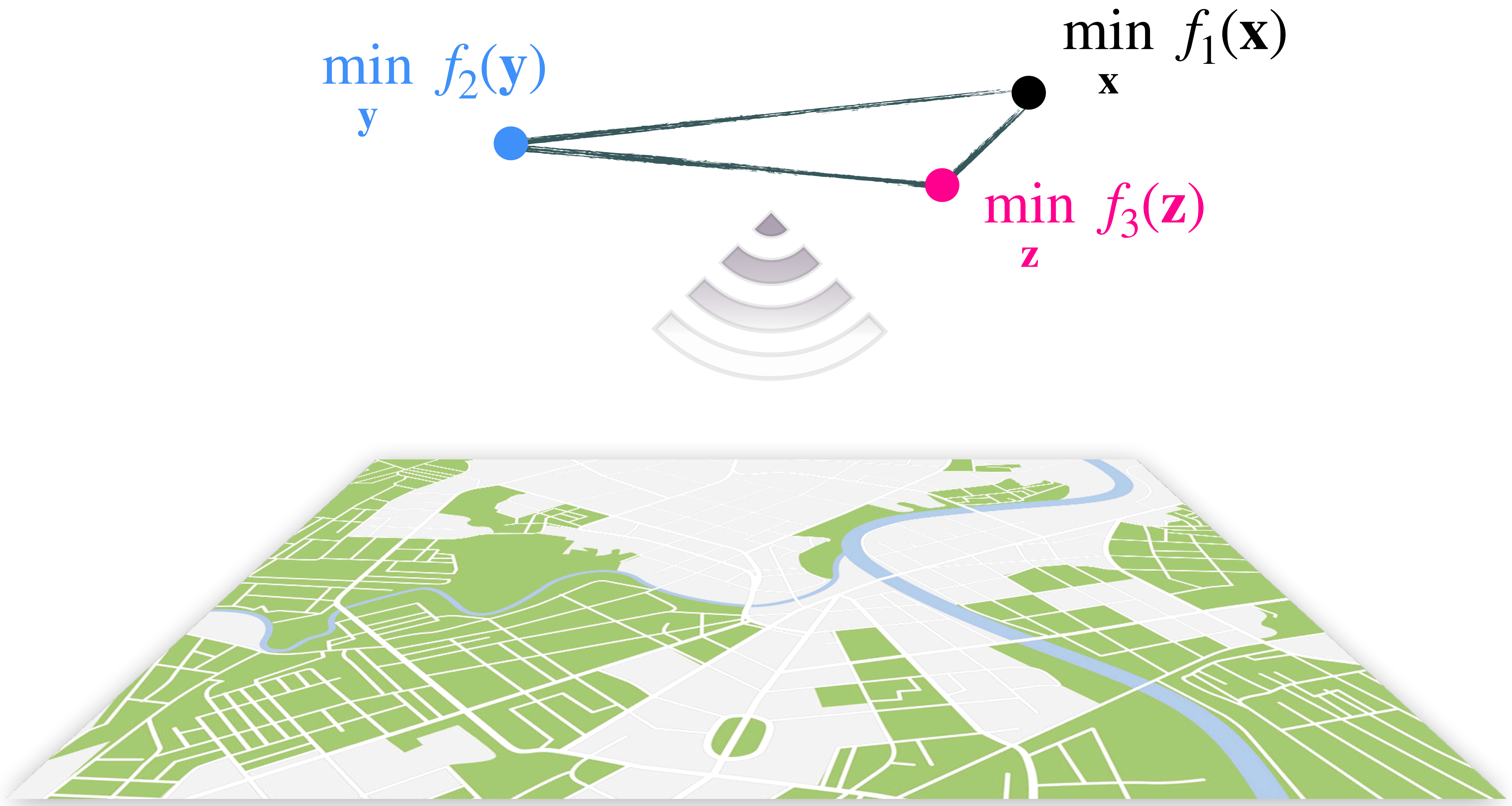
$$x^+ = x - \gamma Df(x)$$

# **Future of autonomy**: optimization problems…

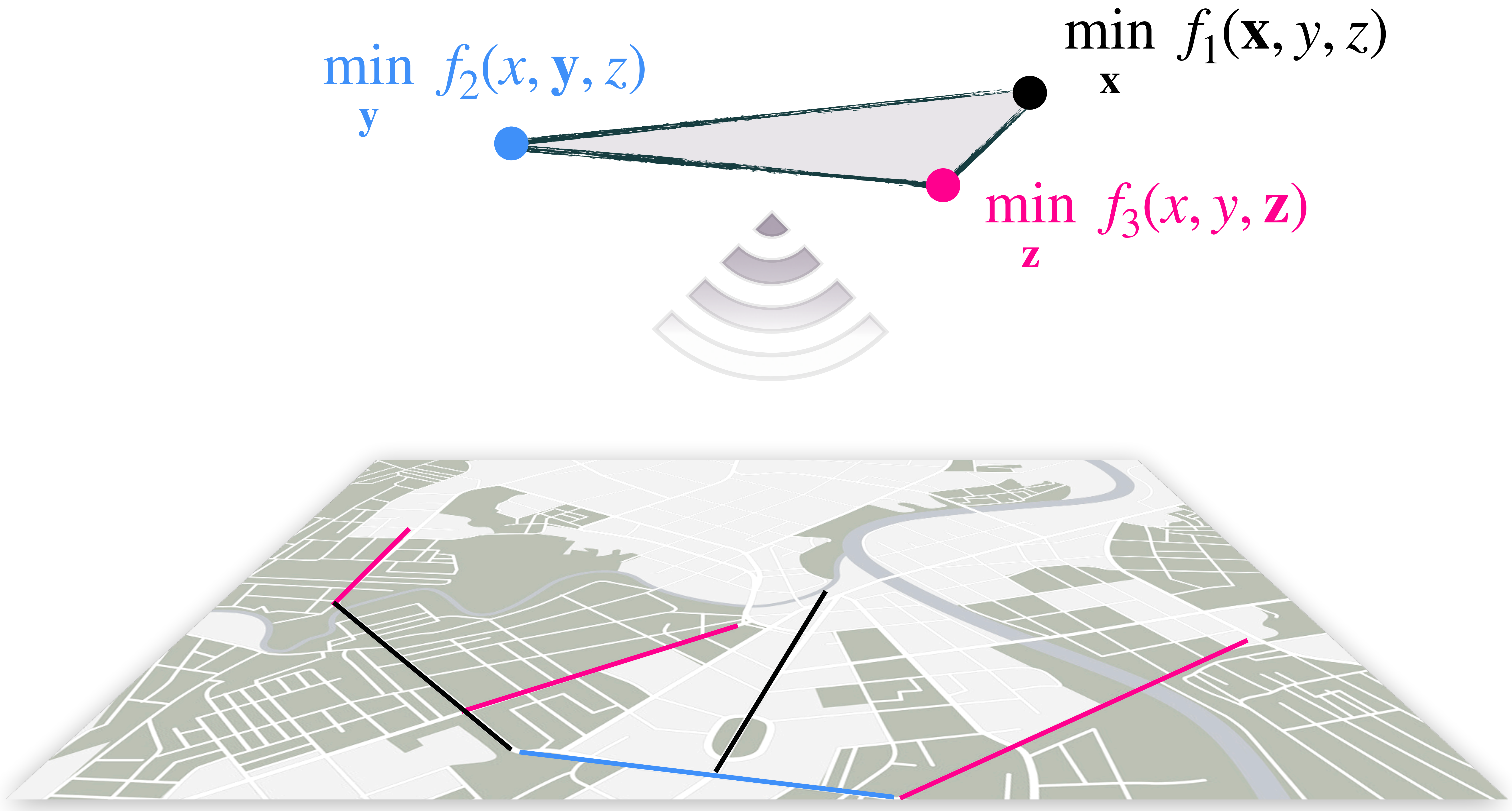


# **Future of autonomy**: optimization problems…

# **Future of autonomy**: optimization problems…



# **Future of autonomy**: *coupled* optimization problems

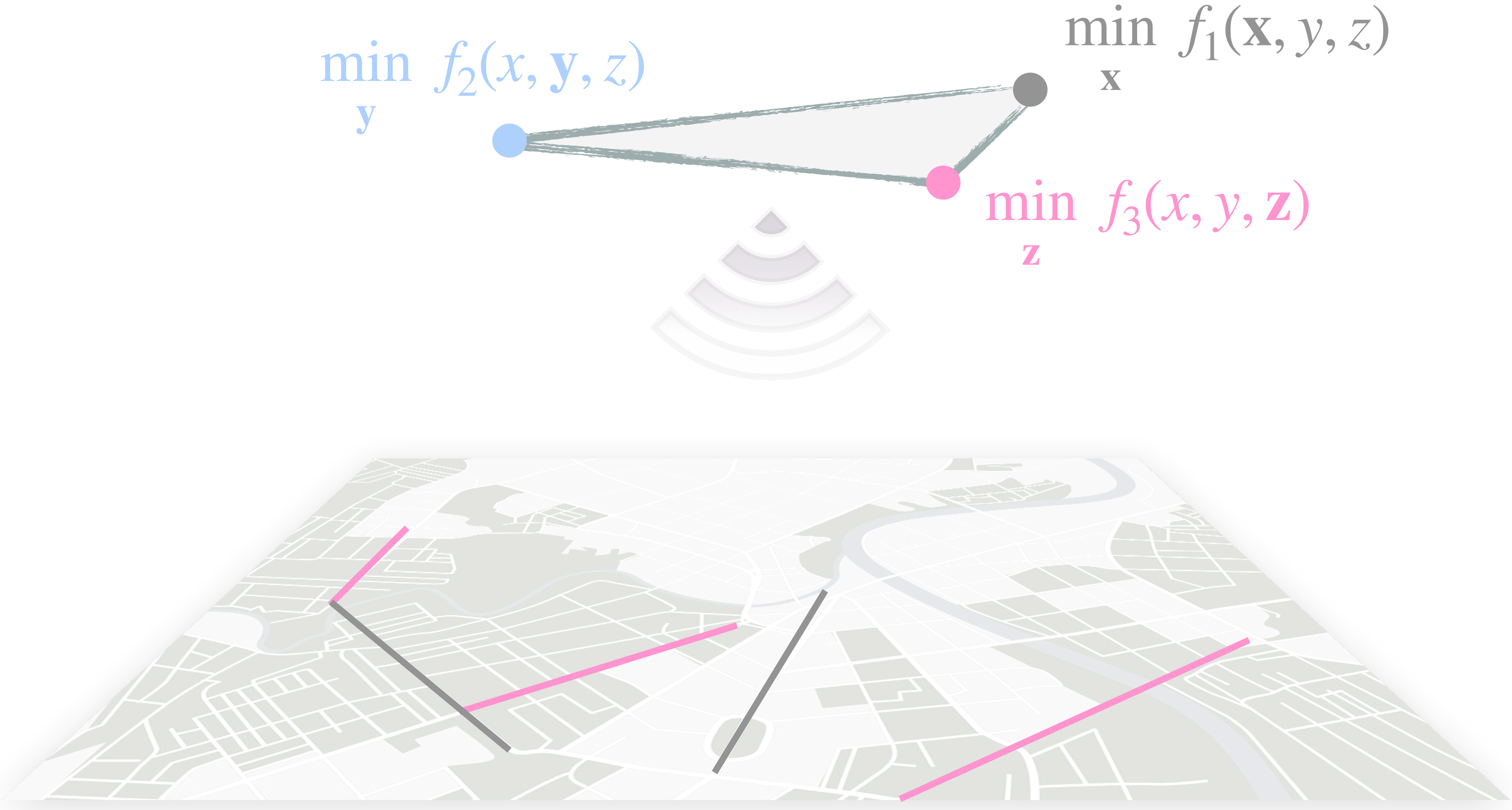


# **Gradient-based learning**: a general framework

$$x^+ = x - \gamma D_1 f_1(\mathbf{x}, y, z)$$

$$y^+ = y - \gamma D_2 f_2(x, \mathbf{y}, z)$$

$$z^+ = z - \gamma D_3 f_3(x, y, \mathbf{z})$$



# **Gradient-based learning**: a general framework

| | Problem Class | Gradient Learning Rule |
|---|---|---|
| | Gradient Play | $x_i^+ = x_i - \gamma_i D_i f_i(x_i, x_{-i})$ |
| ZS | GANs | $\theta^+ = \theta - \gamma \nabla_\theta \ell(\theta, w)$ <br> $w^+ = w + \gamma \nabla_w \ell(\theta, w)$ |
| MARL | Policy Gradient w/ Policy Prediction | $x_1^+ = x_1 - \gamma_1 \nabla_1 J_1(\pi_1(x_1), \pi_2(x_2) + \delta \nabla_2 J_2(\pi_1(x_1), \pi_2(x_2)))$ |
| | Multi-Agent Policy Gradient | $x_i^+ = x_i - \gamma_i \mathbb{E}_{\tau \sim P_\Gamma(\pi)} \left[\sum_t R_i(s_t, u_t) \sum_{k=0}^t \nabla_i \log \pi_i(x_i)(u_{i,k} \mid s_k)\right]$ |
| MAB | Individual Q-Learning | $q_i^+(u_i) = q_i(u_i) + \gamma_i(r_i(u_i, \pi_{-i}(q_i, q_{-i})) - q_i(u_i))$ |
| | Multi-Agent Gradient Bandits | $x_{i,\ell}^+ = x_{i,\ell} + \gamma_i \mathbb{E}[\beta_i R_i(u_i, u_{-i}) \mid u_i = u],\ u \in U_i$ |
| | Multi-Agent Experts | $x_{i,\ell}^+ = x_{i,\ell} + \gamma_i \mathbb{E}[R_i(u_i, u_{-i}) \mid u_i = u],\ u \in U$ |

# **Gradient-based learning**: a general framework

$$x^+ = x - \gamma D_1 f_1(\mathbf{x}, y, z)$$

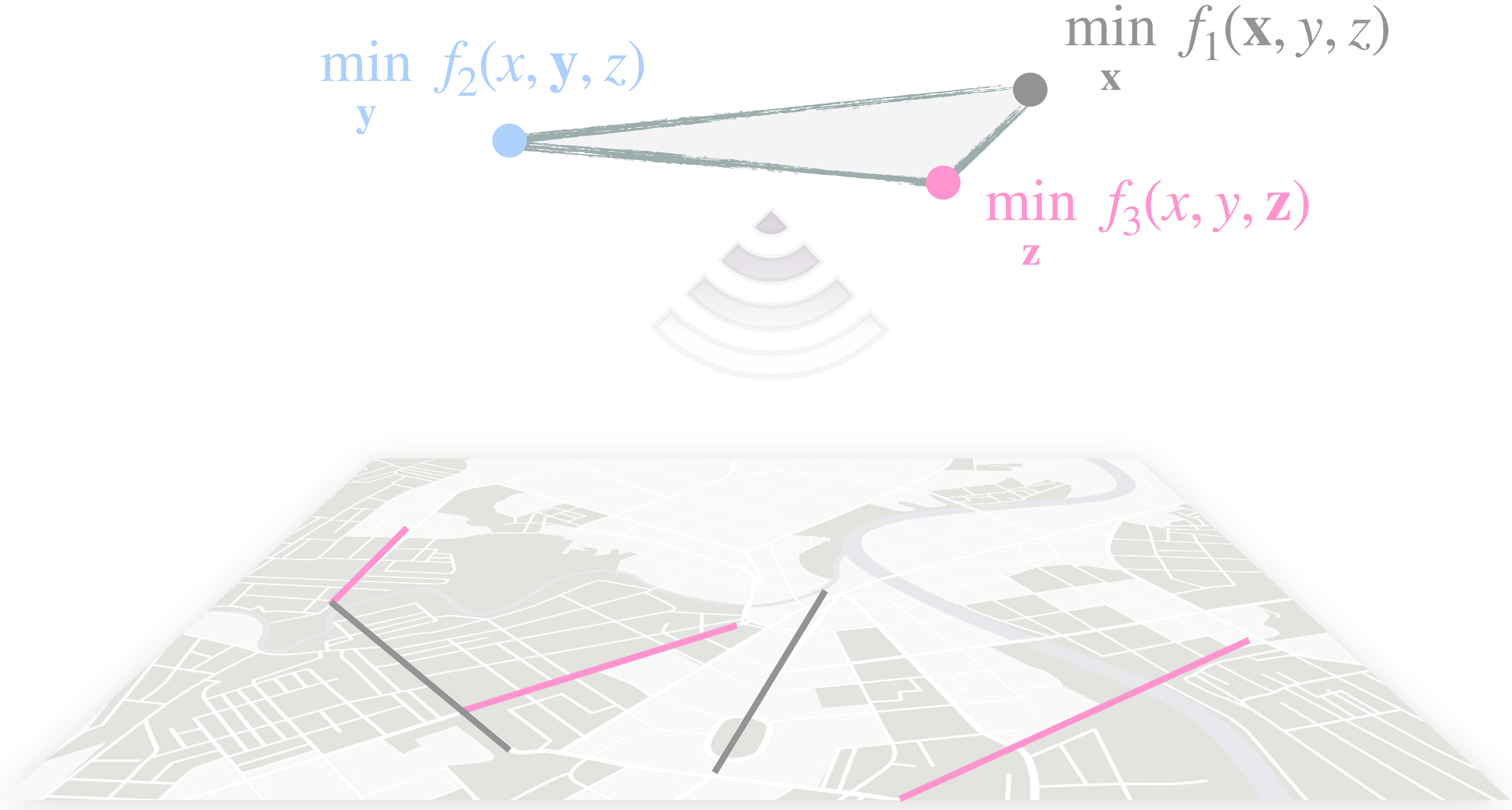
$$y^+ = y - \gamma D_2 f_2(x, \mathbf{y}, z)$$

$$z^+ = z - \gamma D_3 f_3(x, y, \mathbf{z})$$

What does this converge to…

… locally?

… globally?



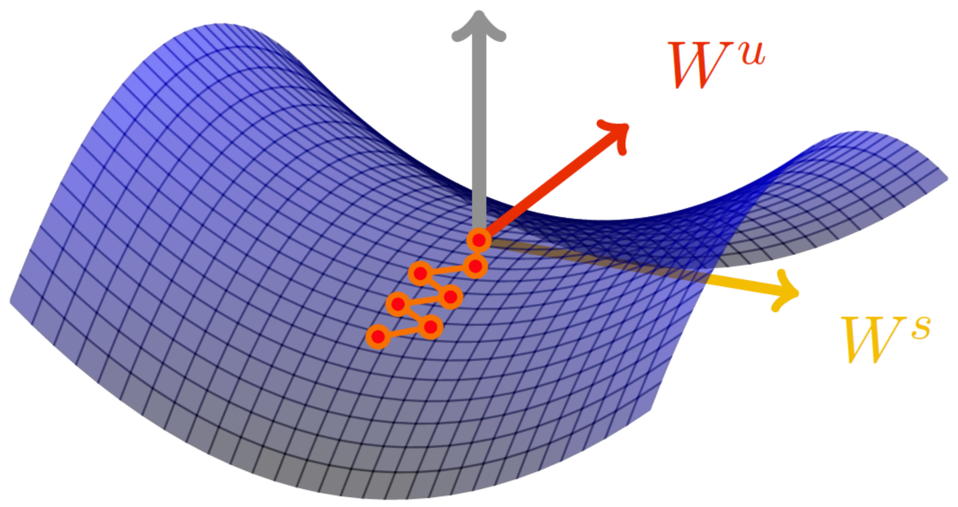
# Overview

- **Part 1:** Local convergence results

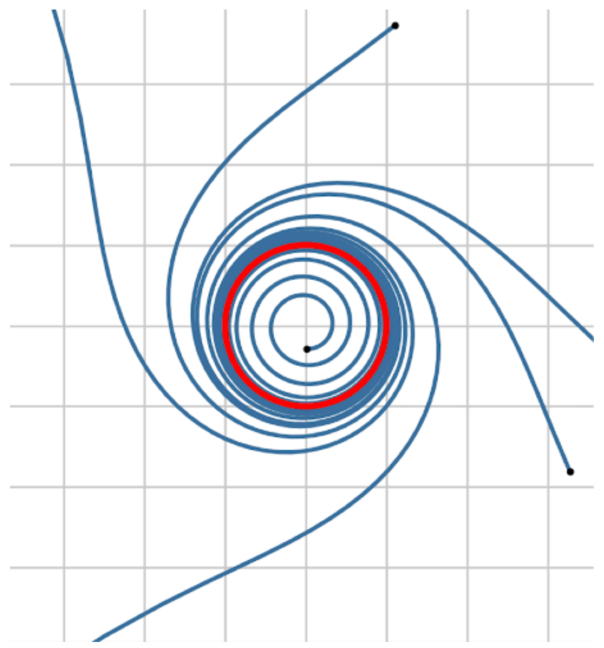
  ✅ Nash equilibria

- **Part 2:** Global limiting behavior of game dynamics

  ⭕ Non-Nash stable equilibria




  ⭕ Periodic orbits



# Overview

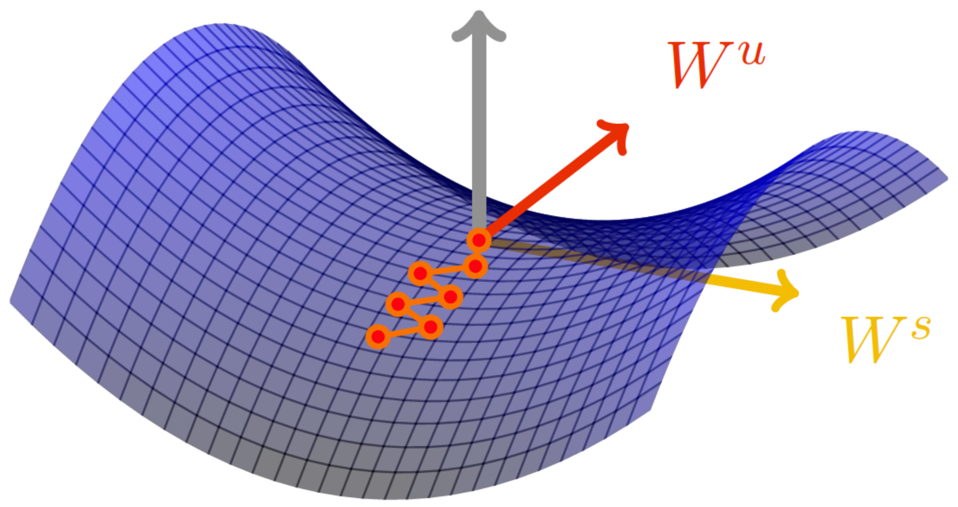
- **Part 1:** Local convergence results

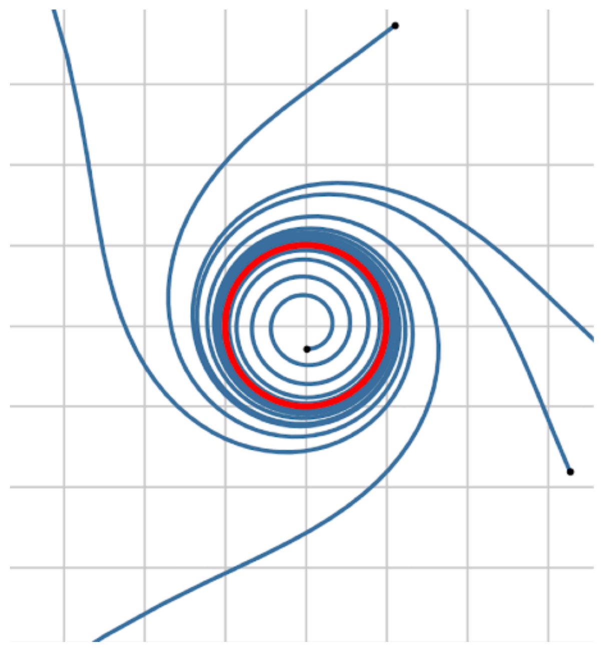
  ✅ Nash equilibria

- **Part 2:** Global limiting behavior of game dynamics

  ⭕ Non-Nash stable equilibria

  ⭕ Periodic orbits



→ Potential Solution!

# Part 1

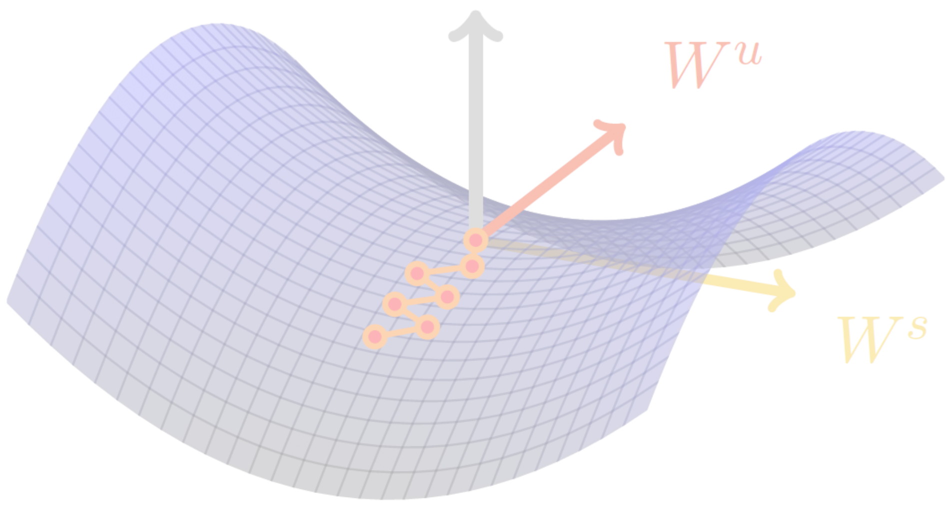
- **Part 1:** Local convergence results

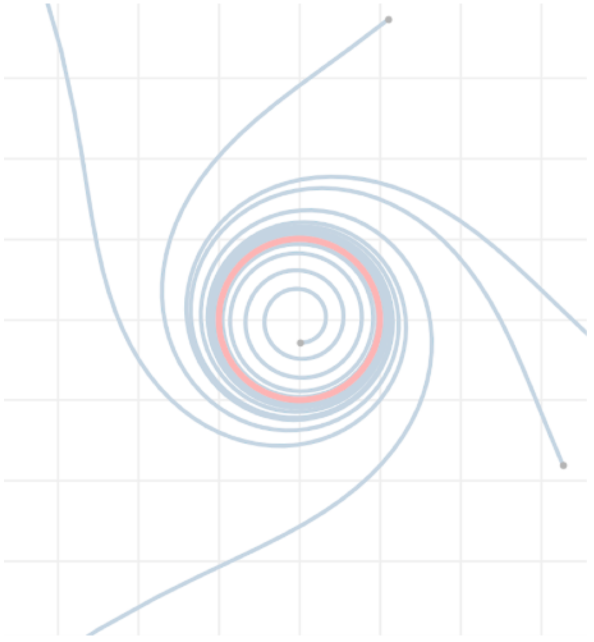
  ✅ Nash equilibria

- **Part 2:** Global limiting behavior of game dynamics

  ⭕ Non-Nash stable equilibria

  ⭕ Periodic orbits



# **Local equilibrium concepts:** Nash equilibria and local minima

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

Ratliff, Burden, Sastry. Genericity and Structural Stability of Non-Degenerate Differential Nash equilibria, ACC 2014

# **Local equilibrium concepts:** Nash equilibria and local minima

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

At a local minimum…

$$Df(x^\star) = 0 \qquad D^2 f(x^\star) \succ 0$$

Ratliff, Burden, Sastry. Genericity and Structural Stability of Non-Degenerate Differential Nash equilibria, ACC 2014

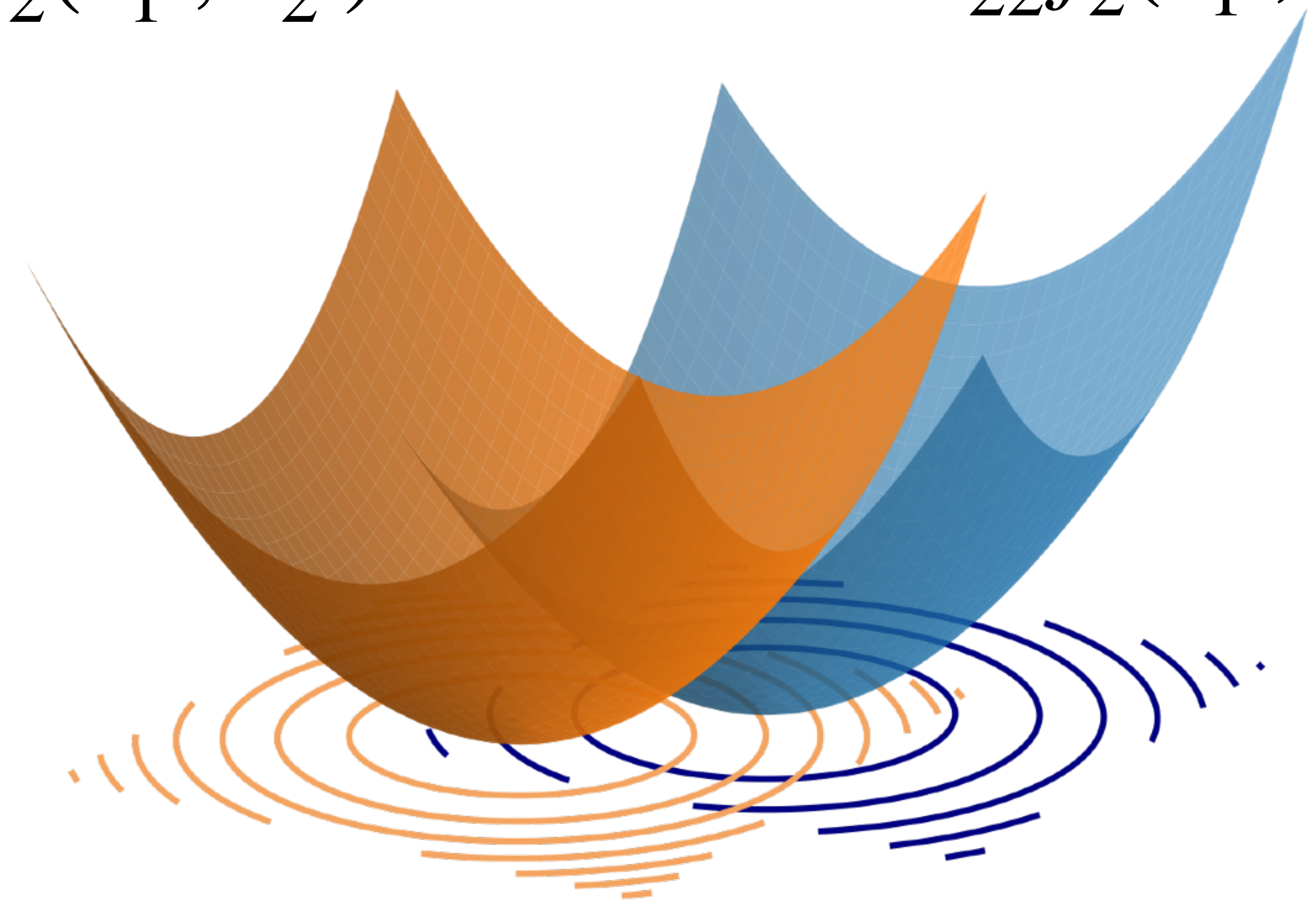
# **Local equilibrium concepts:** Nash equilibria and local minima

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

At a local differential Nash equilibrium…

$$D_1 f_1(x_1^\star, x_2^\star) = 0 \qquad D_{11}^2 f_1(x_1^\star, x_2^\star) \succ 0$$
$$D_2 f_2(x_1^\star, x_2^\star) = 0 \qquad D_{22}^2 f_2(x_1^\star, x_2^\star) \succ 0$$



Gradient descent

$$x^+ = x - \gamma Df(x)$$

At a local minimum…

$$Df(x^\star) = 0 \qquad D^2 f(x^\star) \succ 0$$

Ratliff, Burden, Sastry. Genericity and Structural Stability of Non-Degenerate Differential Nash equilibria, ACC 2014

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

- $f(x)$ decreases at each step

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- $f_i(x_1, x_2)$ can increase

Gradient descent

$$x^+ = x - \gamma Df(x)$$

- $f(x)$ decreases at each step

## **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- $f_i(x_1, x_2)$ can increase

Gradient descent

$$x^+ = x - \gamma Df(x)$$

- $f(x)$ decreases at each step
- $D^2 f(x)$ symmetric Hessian bounded by Lipschitz constant $L$

$$D^2 f(x) = \begin{bmatrix} D_{11}^2 f(x) & D_{12}^2 f(x) \\ D_{12}^2 f(x)^\top & D_{22}^2 f(x) \end{bmatrix} \preceq LI$$

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- $f_i(x_1, x_2)$ can increase
- Non-symmetric game Hessian

$$D\omega(x) = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

- $f(x)$ decreases at each step
- $D^2 f(x)$ symmetric Hessian bounded by Lipschitz constant $L$

$$D^2 f(x) = \begin{bmatrix} D_{11}^2 f(x) & D_{12}^2 f(x) \\ D_{12}^2 f(x)^\top & D_{22}^2 f(x) \end{bmatrix} \preceq LI$$

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- $f_i(x_1, x_2)$ can increase
- Non-symmetric game Hessian

$$D\omega(x) = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

- $f(x)$ decreases at each step
- $D^2 f(x)$ symmetric Hessian bounded by Lipschitz constant $L$

$$D^2 f(x) = \begin{bmatrix} D_{11}^2 f(x) & D_{12}^2 f(x) \\ D_{12}^2 f(x)^\top & D_{22}^2 f(x) \end{bmatrix} \preceq LI$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D \begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

**Example 1**: Potential game ( $\exists \phi : \; D_i f_i \equiv D_i \phi$ )

$$D\omega = \begin{bmatrix} D_{11}^2 \phi(x) & D_{12}^2 \phi(x) \\ D_{12}^2 \phi(x) & D_{22}^2 \phi(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ S_2^\top & S_3 \end{bmatrix}$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D^2_{11} f_1(x) & D^2_{12} f_1(x) \\ D^2_{21} f_2(x) & D^2_{22} f_2(x) \end{bmatrix}$$

**Example 1**: Potential game ( $\exists \phi : \ D_i f_i \equiv D_i \phi$ )

$$D\omega = \begin{bmatrix} D^2_{11}\phi(x) & D^2_{12}\phi(x) \\ D^2_{12}\phi(x) & D^2_{22}\phi(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ S_2^\top & S_3 \end{bmatrix}$$

**Symmetric hessian**:
can be solved using
optimization!

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Example 1: Potential game ( $\exists \phi : \; D_i f_i \equiv D_i \phi$ )
**Example 2**: Zero-sum game ( $f_1 = -f_2 = f$ )

$$D\omega = \begin{bmatrix} D_{11}^2 f(x) & D_{21}^2 f(x) \\ -D_{12}^2 f(x) & -D_{22}^2 f(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ -S_2^\top & S_3 \end{bmatrix}$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D \begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Example 1: Potential game ( $\exists \phi : \; D_i f_i \equiv D_i \phi$ )
**Example 2**: Zero-sum game ( $f_1 = -f_2 = f$ )

$$D\omega = \begin{bmatrix} D_{11}^2 f(x) & D_{21}^2 f(x) \\ -D_{12}^2 f(x) & -D_{22}^2 f(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ -S_2^\top & S_3 \end{bmatrix}$$

**Anti-symmetric off diagonals:**
Some ways to solve (LPs… etc)

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Example 1: Potential game ( $\exists \phi : \ D_i f_i \equiv D_i \phi$ )
Example 2: Zero-sum game ( $f_1 = -f_2 = f$ )
**Example 3**: General-sum non-cooperative game

$$D\omega = \begin{bmatrix} D_{11}^2 f_1(x) & D_{21}^2 f_1(x) \\ D_{12}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}$$

# **Examples of games:** potential, zero-sum, general-sum

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

- **Definition**: Game Hessian

$$D\omega \triangleq D\begin{bmatrix} D_1 f_1(x) \\ D_2 f_2(x) \end{bmatrix} = \begin{bmatrix} D_{11}^2 f_1(x) & D_{12}^2 f_1(x) \\ D_{21}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix}$$

Example 1: Potential game ( $\exists \phi : \; D_i f_i \equiv D_i \phi$ )
Example 2: Zero-sum game ( $f_1 = -f_2 = f$ )
**Example 3**: General-sum non-cooperative game

$$D\omega = \begin{bmatrix} D_{11}^2 f_1(x) & D_{21}^2 f_1(x) \\ D_{12}^2 f_2(x) & D_{22}^2 f_2(x) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}$$

**Difficult to solve in general!**

# **Local convergence analysis:** gradient-play vs. gradient descent

Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$

$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

Gradient descent

$$x^+ = x - \gamma Df(x)$$

**Classical result:**

$\mu$-strongly convex and $L$-smooth

$$\mu \preceq D^2 f(x) \preceq L\,.$$

With learning rate $\gamma = 1/L$

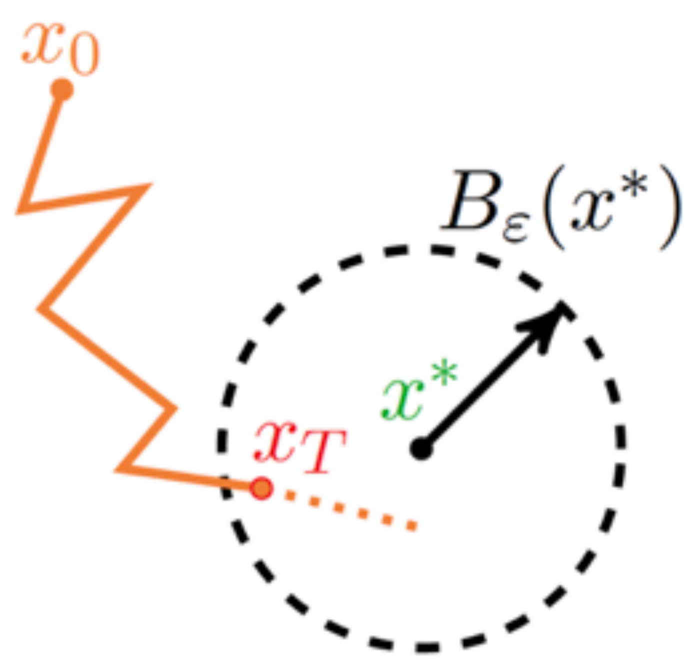
$x^{(T)}$ approaches $x^\star$ in $T$ iterations:

$$\boxed{\|x^{(T)} - x^\star\| \leq \exp\left(-\frac{\mu}{L}T\right)\|x^{(1)} - x^\star\|}$$

# **Local convergence analysis:** gradient-play vs. gradient descent

## Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$




## Gradient descent

$$x^+ = x - \gamma Df(x)$$

**Classical result:**

$\mu$-strongly convex and $L$-smooth

$$\mu \preceq D^2 f(x) \preceq L\,.$$

With learning rate $\gamma = 1/L$

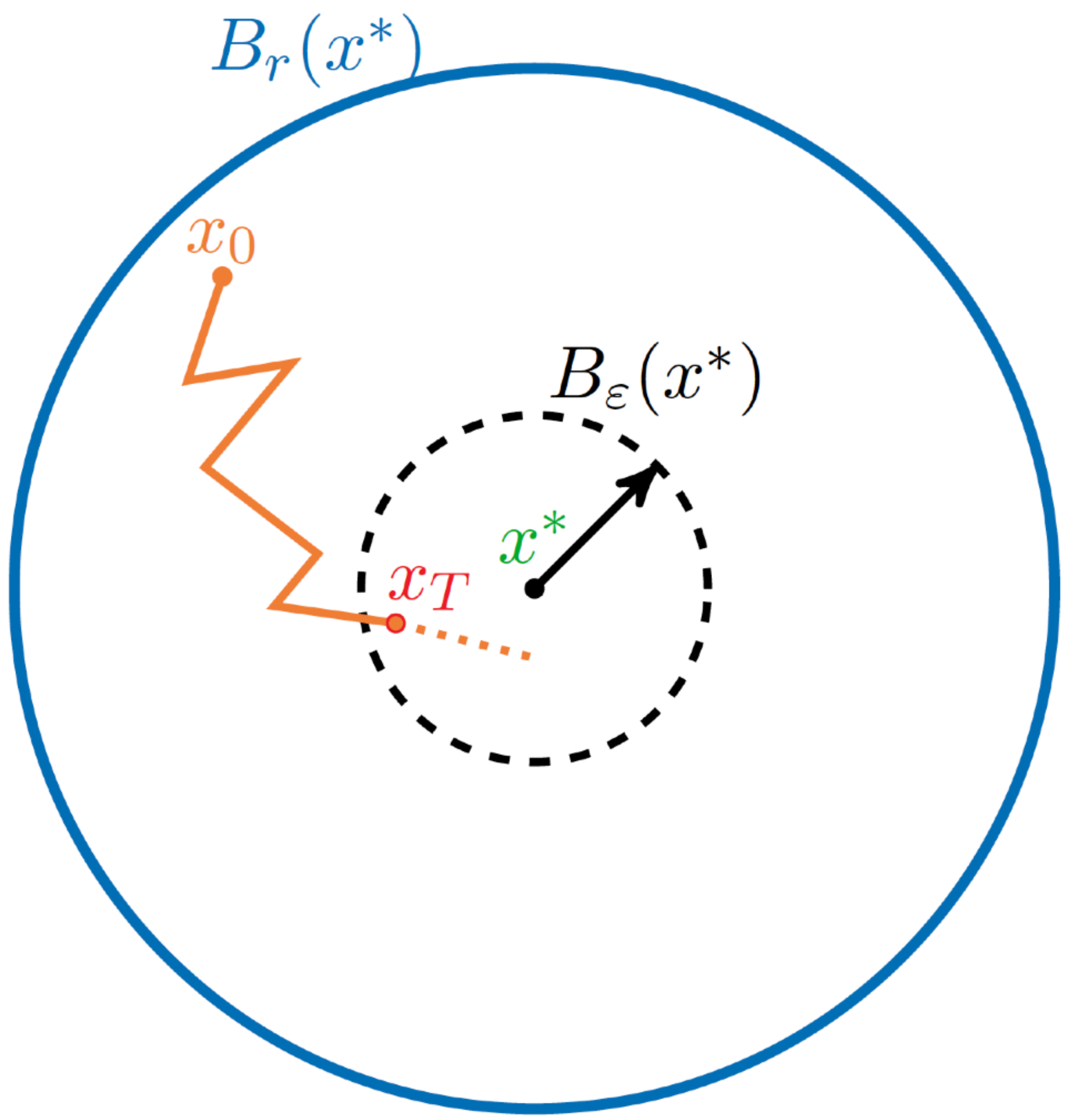
$x^{(T)}$ approaches $x^\star$ in $T$ iterations:

$$\boxed{\|x^{(T)} - x^\star\| \le \exp\left(-\frac{\mu}{L}T\right)\|x^{(1)} - x^\star\|}$$

# **Local convergence analysis:** gradient-play vs. gradient descent

## Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$




## Gradient descent

$$x^+ = x - \gamma Df(x)$$

**Classical result:**

$\mu$-strongly convex and $L$-smooth

$$\mu \preceq D^2 f(x) \preceq L\,.$$

With learning rate $\gamma = 1/L$

$x^{(T)}$ approaches $x^\star$ in $T$ iterations:

$$\boxed{\|x^{(T)} - x^\star\| \leq \exp\left(-\frac{\mu}{L}T\right) \|x^{(1)} - x^\star\|}$$

# **Local convergence analysis:** gradient-play vs. gradient descent

## Gradient-play

$$x_1^+ = x_1 - \gamma D_1 f_1(x_1, x_2)$$
$$x_2^+ = x_2 - \gamma D_2 f_2(x_1, x_2)$$

**Main theorem** (informal):

$$\alpha = \min_{x \in B_r(x)} \sigma_{\min} \overbrace{(D\omega(x)^\top + D\omega(x))/2}^{\text{symmetric part of } D\omega}$$

$$\beta = \max_{x \in B_r(x)} \sigma_{\max} \left(D\omega(x)\right)$$

With learning rate $\gamma = \alpha/\beta^2 \ldots .$

$$\boxed{\|x^{(T)} - x^\star\| \le \exp\left(-\frac{\alpha^2}{2\beta^2} T\right) \|x^{(1)} - x^\star\|}$$

## Gradient descent

$$x^+ = x - \gamma Df(x)$$

**Classical result:**

$\mu$-strongly convex and $L$-smooth

$$\mu \preceq D^2 f(x) \preceq L\,.$$

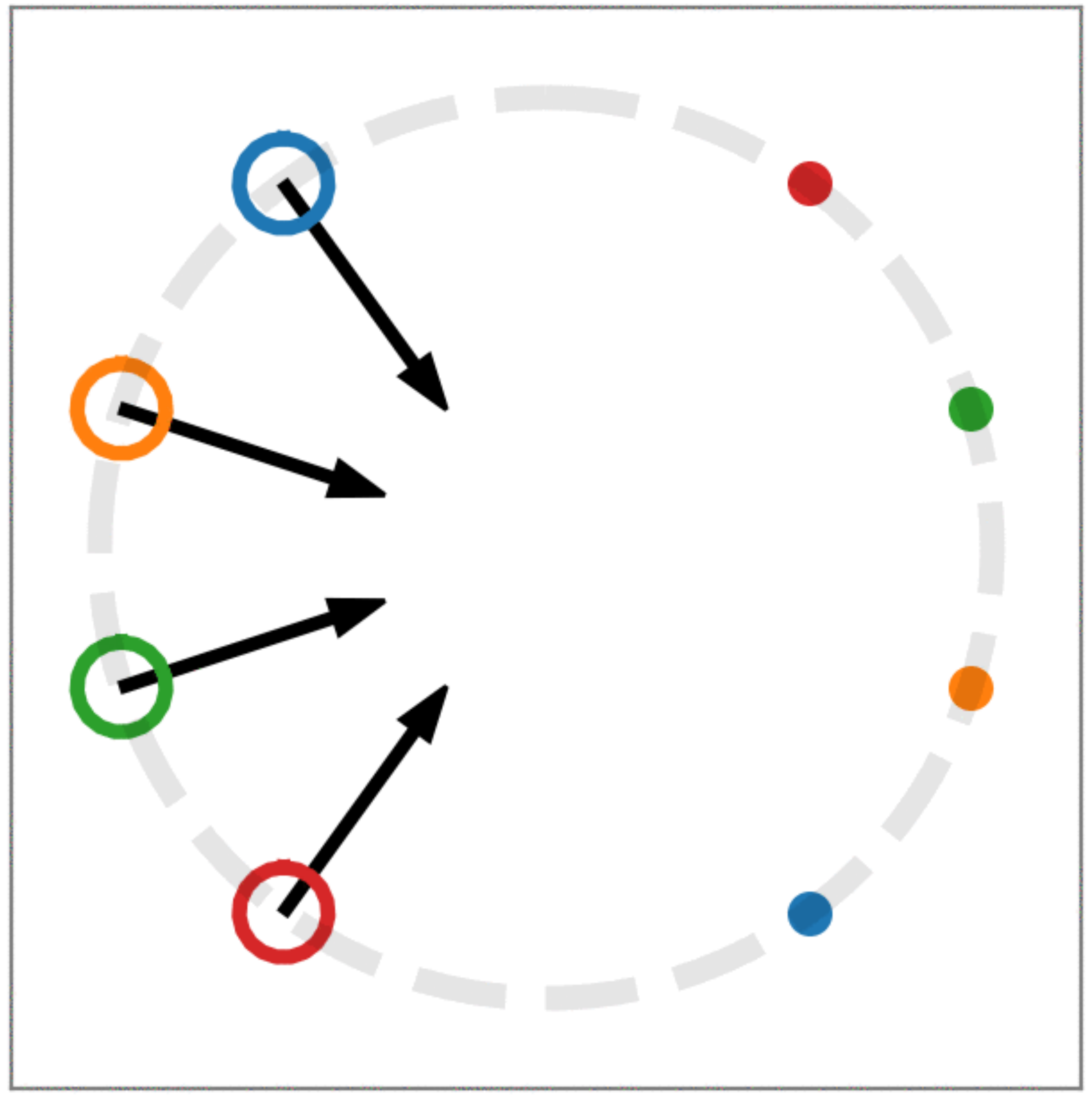
With learning rate $\gamma = 1/L$

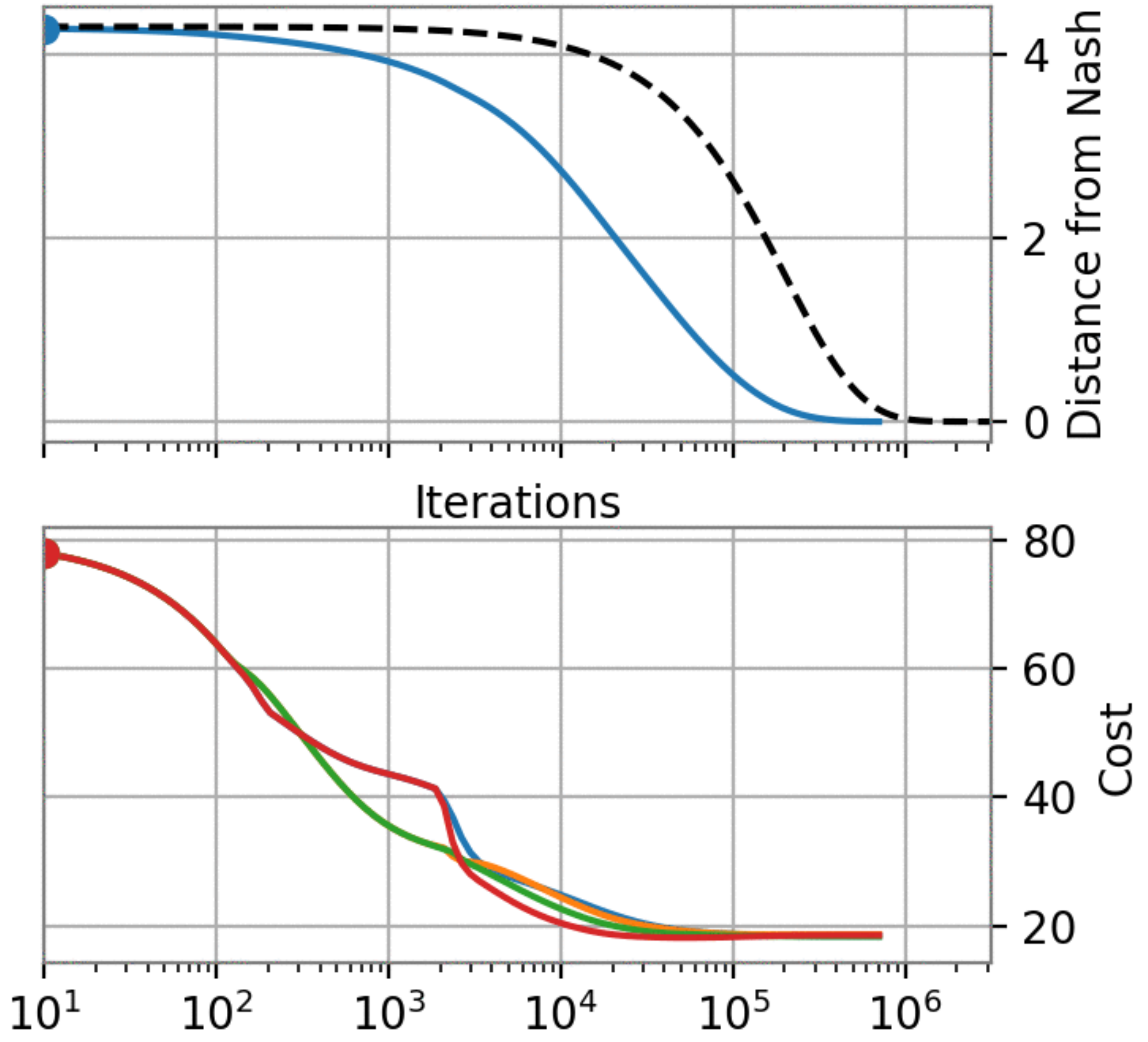
$x^{(T)}$ approaches $x^\star$ in $T$ iterations:

$$\boxed{\|x^{(T)} - x^\star\| \le \exp\left(-\frac{\mu}{L} T\right) \|x^{(1)} - x^\star\|}$$

L.J. Ratliff, B. Chasnov, D. Calderone, E. Mazumdar, S. Burden. Convergence Guarantees for Gradient-Based Learning in Continuous Games, under review 2018.

# **Dynamic game**: convergence to Nash equilibria






$$u_1^+ = u_1 - \gamma D_1 f_1(\mathbf{u})$$
$$u_2^+ = u_2 - \gamma D_2 f_2(\mathbf{u})$$
$$u_3^+ = u_3 - \gamma D_3 f_3(\mathbf{u})$$
$$u_4^+ = u_4 - \gamma D_4 f_4(\mathbf{u})$$

$$u_i = \begin{bmatrix} \mathbf{u_i}(1) \\ \vdots \\ \mathbf{u}_i(k) \\ \vdots \\ \mathbf{u}_i(N) \end{bmatrix}$$

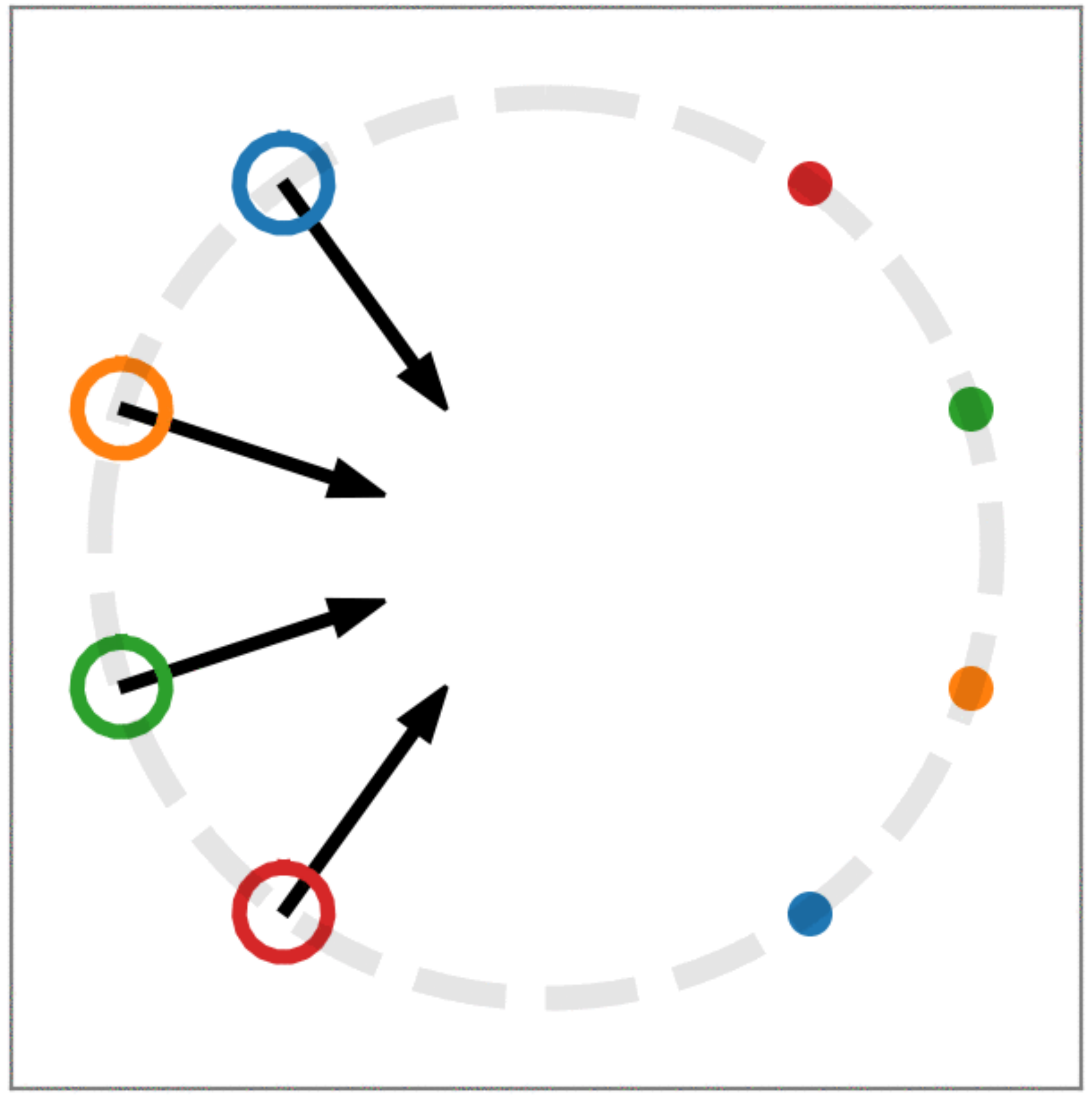
Linear dynamics

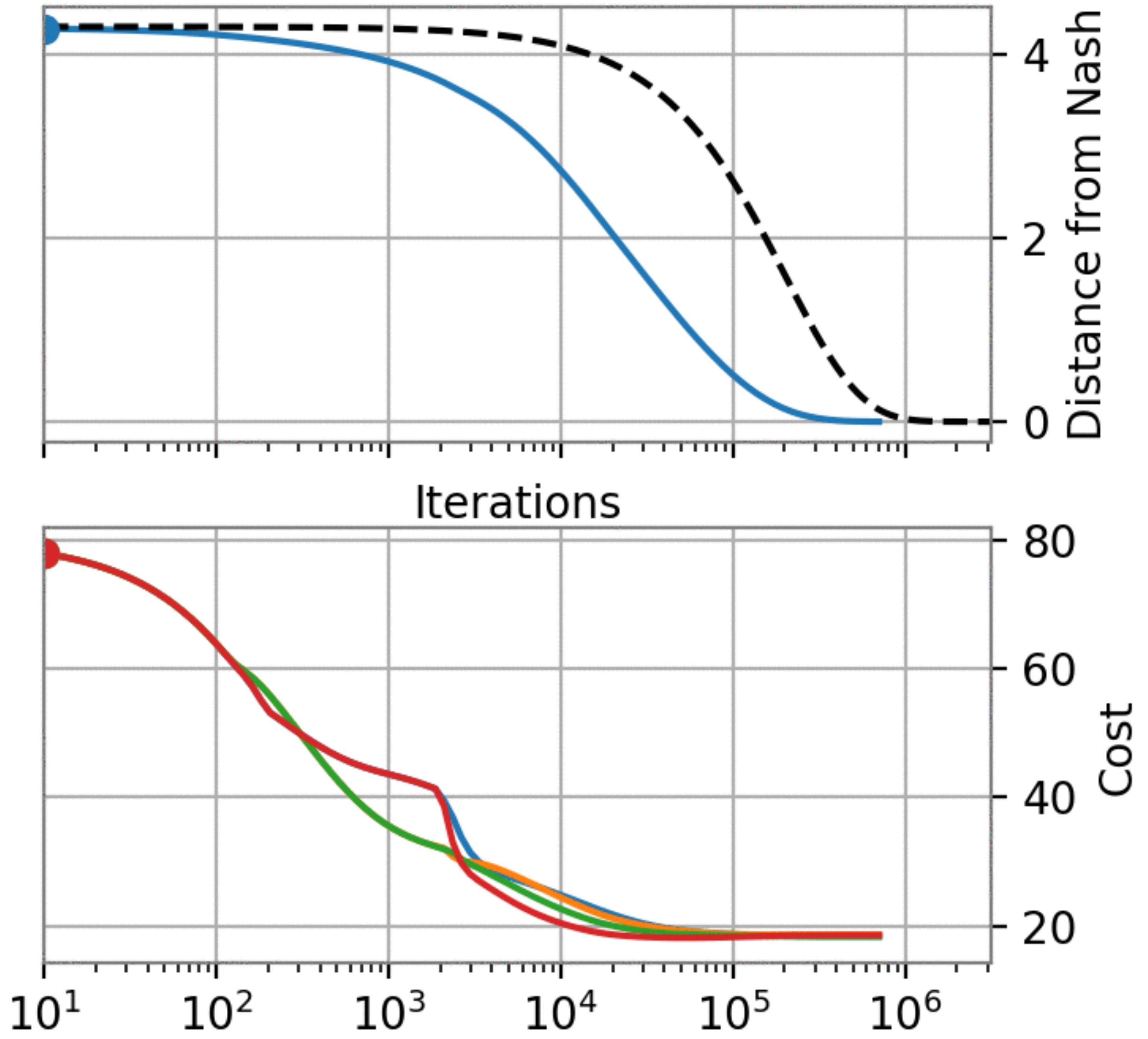
$$z_i(k+1) = Az_i(k) + B\mathbf{u_i}(k)$$

General-sum costs

$$f_i(u_1, \cdots, u_4) = (\text{go to destination})^2 + \exp(\text{avoid each other})^2$$

# **Dynamic game**: convergence to Nash equilibria






$$u_1^+ = u_1 - \gamma D_1 f_1(\mathbf{u})$$
$$u_2^+ = u_2 - \gamma D_2 f_2(\mathbf{u})$$
$$u_3^+ = u_3 - \gamma D_3 f_3(\mathbf{u})$$
$$u_4^+ = u_4 - \gamma D_4 f_4(\mathbf{u})$$

Linear dynamics

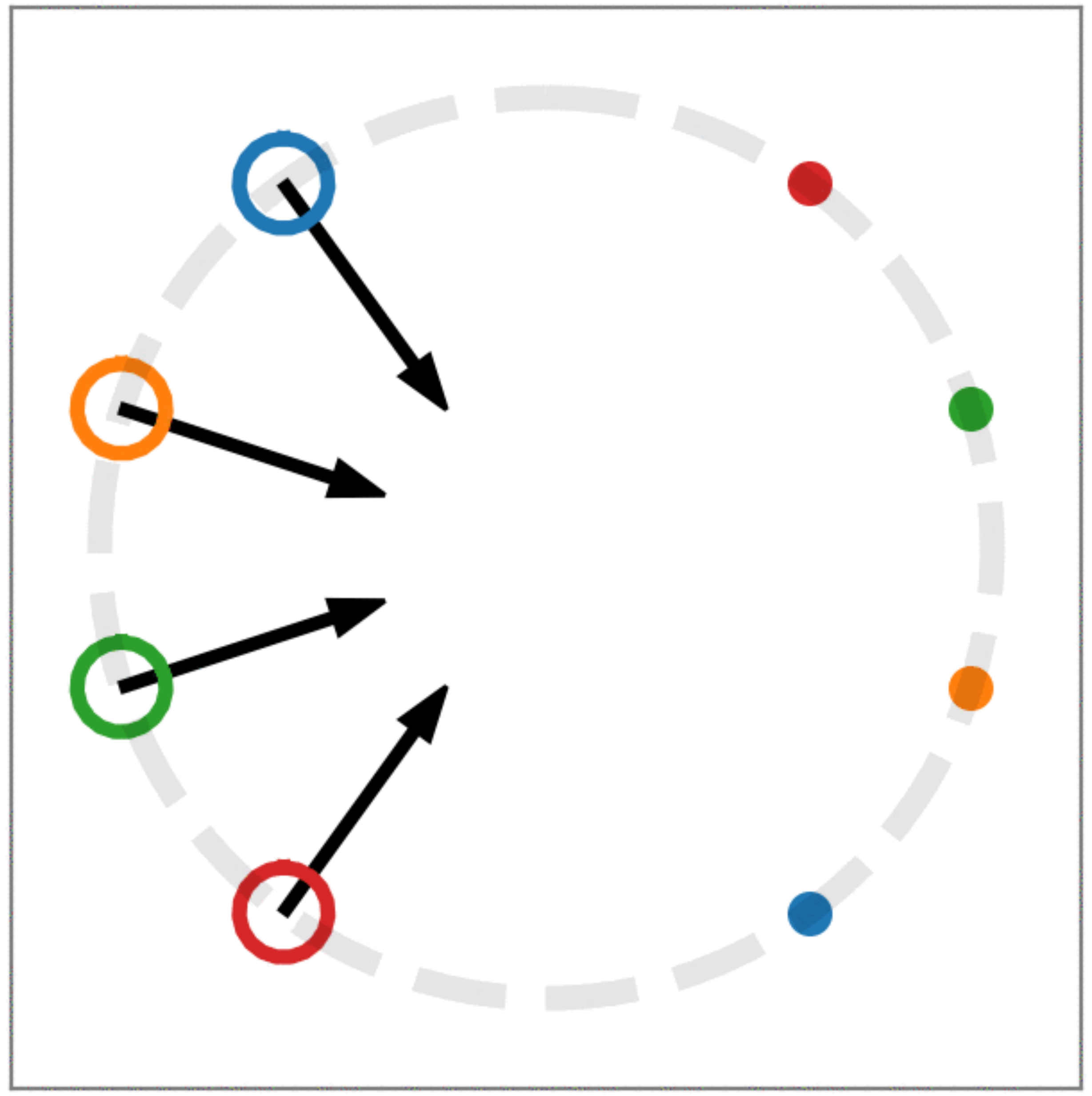
$$z_i(k+1) = Az_i(k) + B\mathbf{u_i}(k)$$

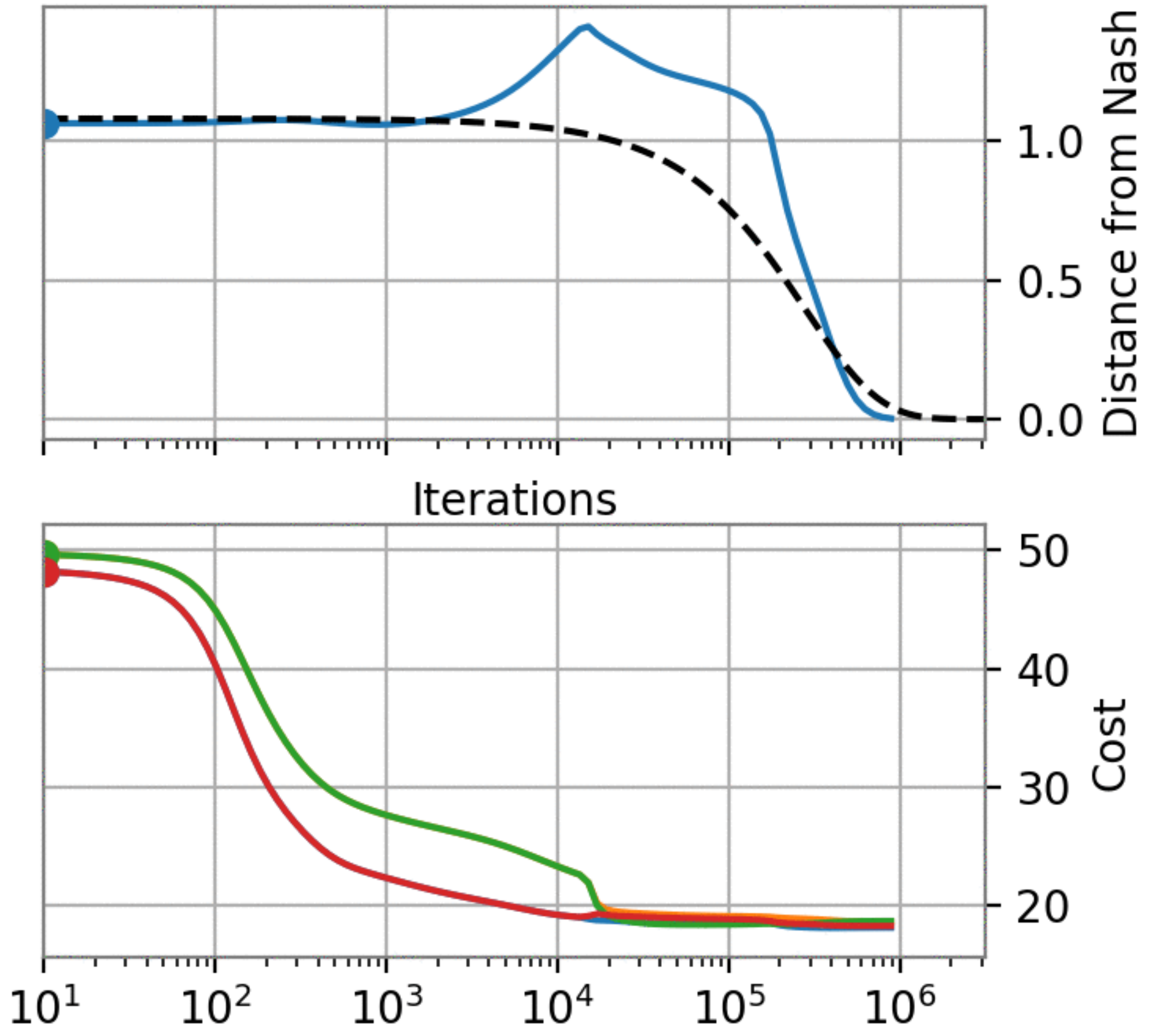
General-sum costs

$$f_i(u_1, \cdots, u_4) = (\text{go to destination})^2 + \exp(\text{avoid each other})^2$$

$$u_i = \begin{bmatrix} \mathbf{u_i}(1) \\ \vdots \\ \mathbf{u}_i(k) \\ \vdots \\ \mathbf{u}_i(N) \end{bmatrix}$$

# **Dynamic game**: convergence to Nash equilibria






$$u_1^+ = u_1 - \gamma D_1 f_1(\mathbf{u})$$
$$u_2^+ = u_2 - \gamma D_2 f_2(\mathbf{u})$$
$$u_3^+ = u_3 - \gamma D_3 f_3(\mathbf{u})$$
$$u_4^+ = u_4 - \gamma D_4 f_4(\mathbf{u})$$

Linear dynamics

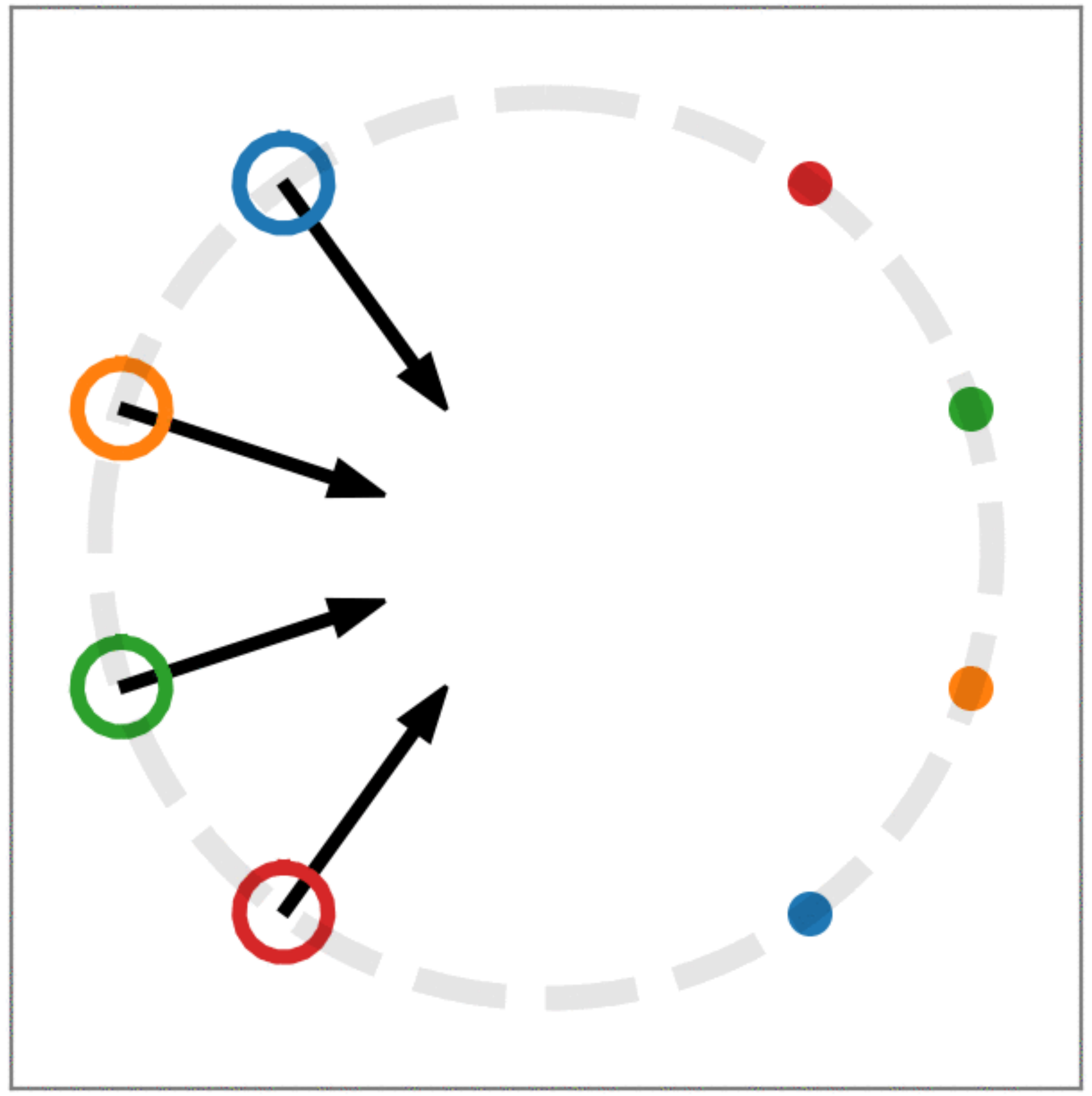
$$z_i(k+1) = Az_i(k) + B\mathbf{u_i}(k)$$

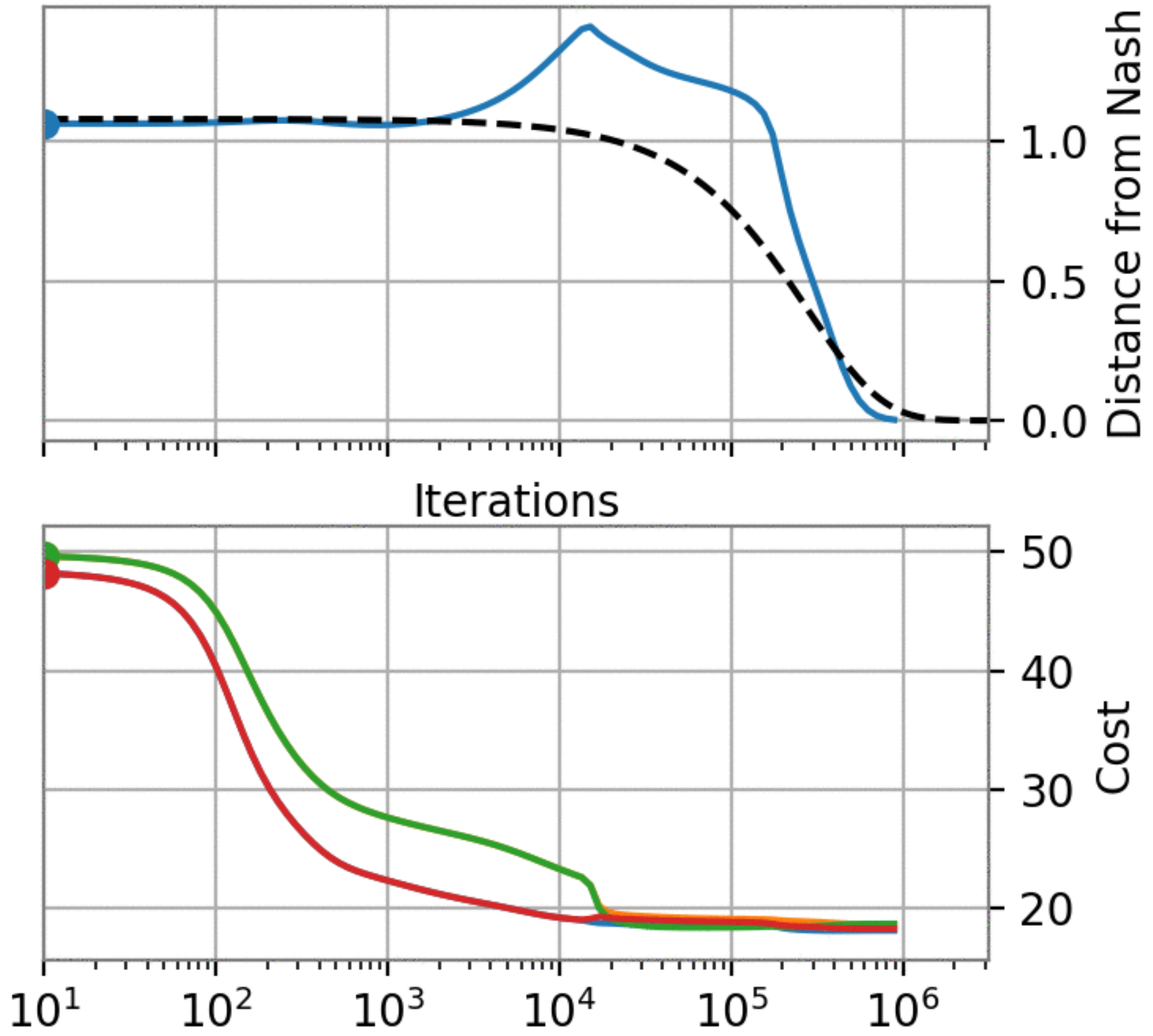
General-sum costs

$$f_i(u_1, \cdots, u_4) = (\text{go to destination})^2 + \exp(\text{avoid each other})^2$$

$$u_i = \begin{bmatrix} \mathbf{u_i}(1) \\ \vdots \\ \mathbf{u}_i(k) \\ \vdots \\ \mathbf{u}_i(N) \end{bmatrix}$$

# **Dynamic game**: convergence to Nash equilibria






$$u_1^+ = u_1 - \gamma D_1 f_1(\mathbf{u})$$
$$u_2^+ = u_2 - \gamma D_2 f_2(\mathbf{u})$$
$$u_3^+ = u_3 - \gamma D_3 f_3(\mathbf{u})$$
$$u_4^+ = u_4 - \gamma D_4 f_4(\mathbf{u})$$

Linear dynamics

$$z_i(k+1) = A z_i(k) + B\mathbf{u_i}(k)$$

General-sum costs

$$f_i(u_1, \cdots, u_4) = (\text{go to destination})^2 + \exp(\text{avoid each other})^2$$

$$u_i = \begin{bmatrix} \mathbf{u_i}(1) \\ \vdots \\ \mathbf{u}_i(k) \\ \vdots \\ \mathbf{u}_i(N) \end{bmatrix}$$

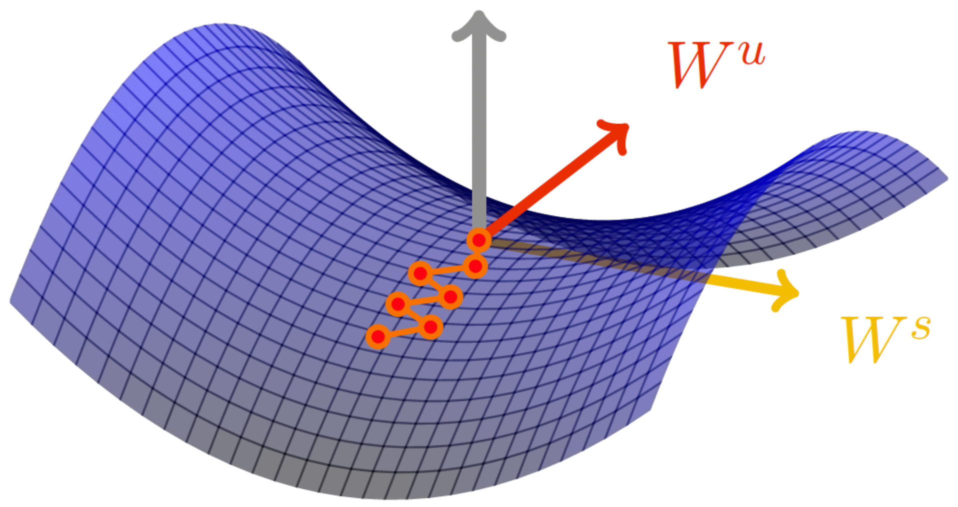
# Part 2: What can go wrong?

- **Part 1:** Local convergence results

  ✅ Nash equilibria

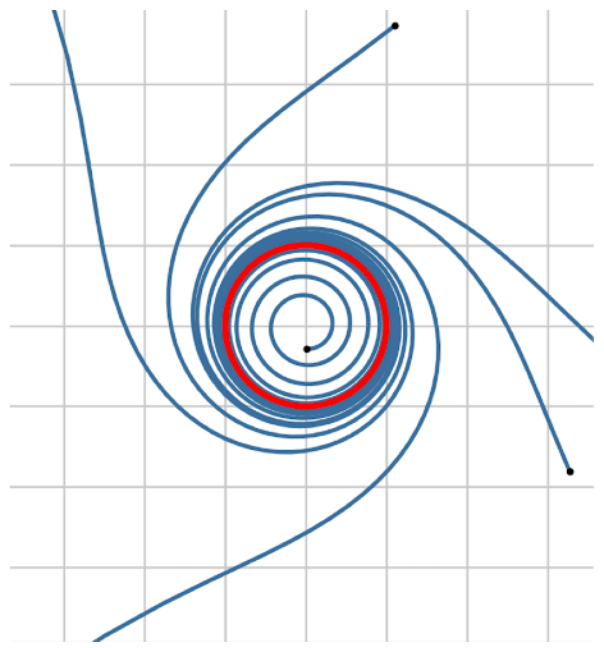
- **Part 2:** Global limiting behavior of game dynamics

  ⭕ Non-Nash stable equilibria





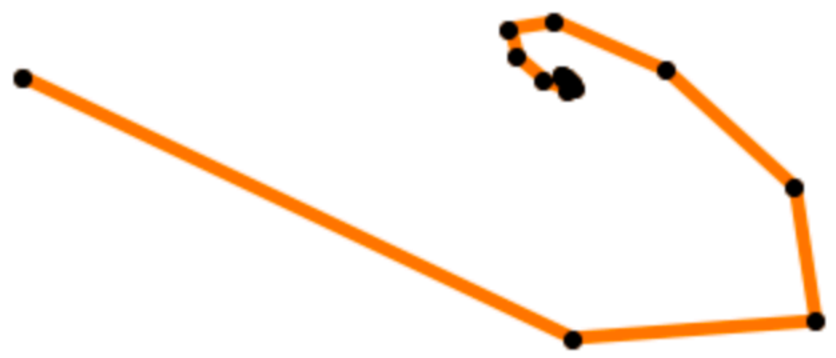
  ⭕ Periodic orbits



# **Part 2: What can go wrong?** Limiting dynamics of gradient-play

Gradient-play

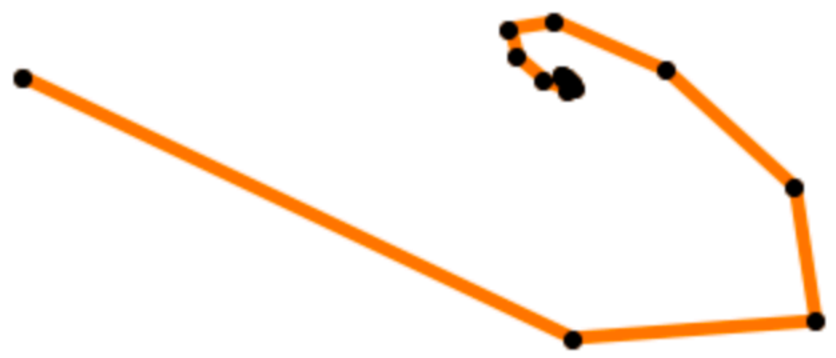
$$x_i^+ \leftarrow x_i - \gamma D_i f_i(x_i, x_{-i})$$



# **Part 2: What can go wrong?** Limiting dynamics of gradient-play

Gradient-play

$$x_i^+ \leftarrow x_i - \gamma D_i f_i(x_i, x_{-i})$$

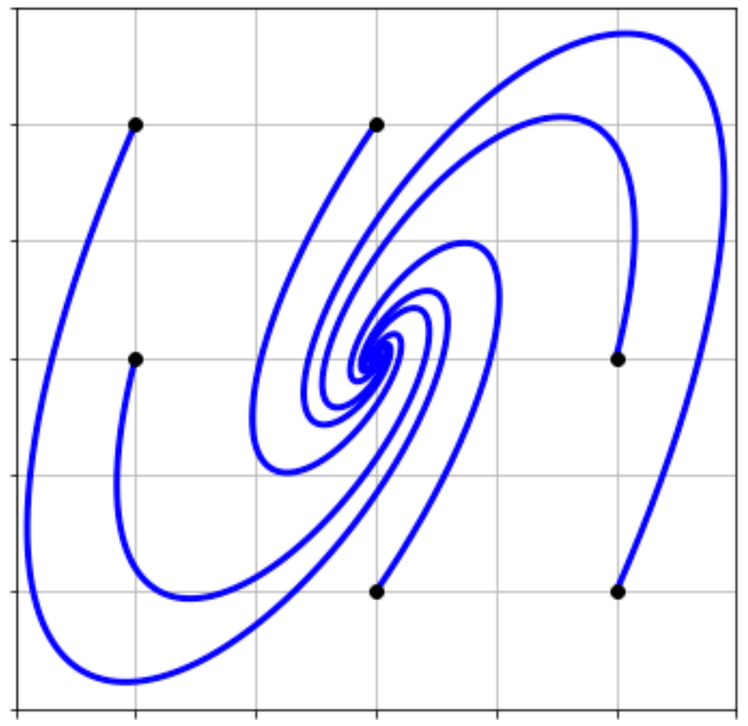


$\updownarrow$

Continuous time dynamical system

$$\dot{x} = -\omega(x)$$

# **Non-Nash stable equilibria:** saddle point

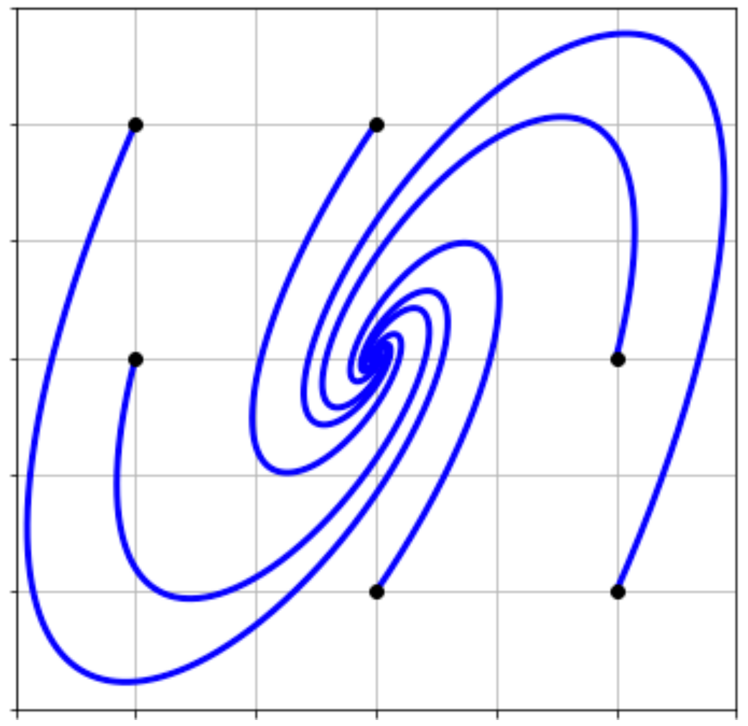


$$\dot{x} = -\,\omega(x)$$

$$D\omega = \begin{bmatrix} - & \\ & + \end{bmatrix}, \quad \text{spec}(D\omega) \subset \mathbb{C}_+^\circ$$

Example:

## **Non-Nash stable equilibria:** saddle point



$\dot{x} = -\omega(x)$

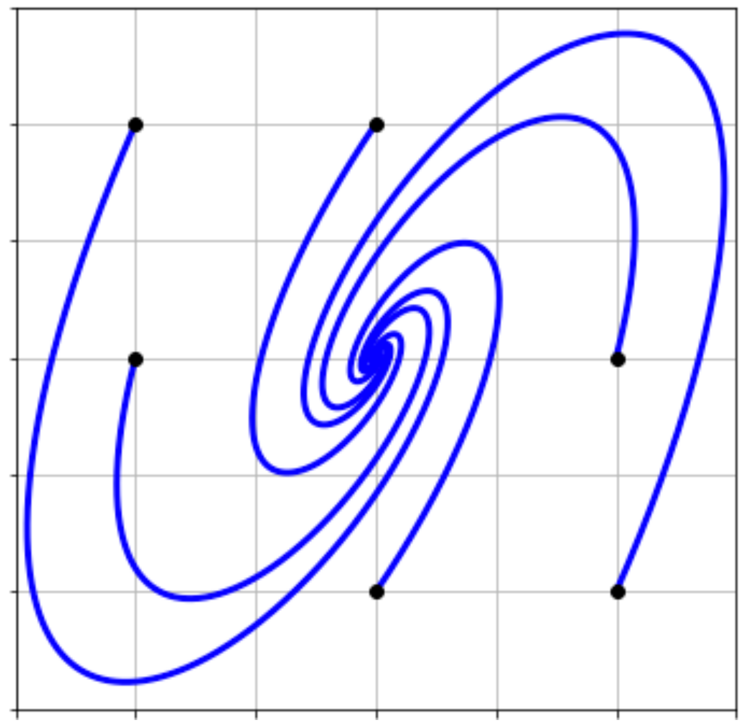
$$D\omega = \begin{bmatrix} - & \\ & + \end{bmatrix}, \quad \text{spec}(D\omega) \subset \mathbb{C}_+^\circ$$

Example:

$$f_1(x_1, x_2) = -x_1^2 + 4x_1x_2$$

$$f_2(x_1, x_2) = 6x_2^2 - 8x_1x_2$$

# **Non-Nash stable equilibria:** saddle point



$\dot{x} = -\omega(x)$

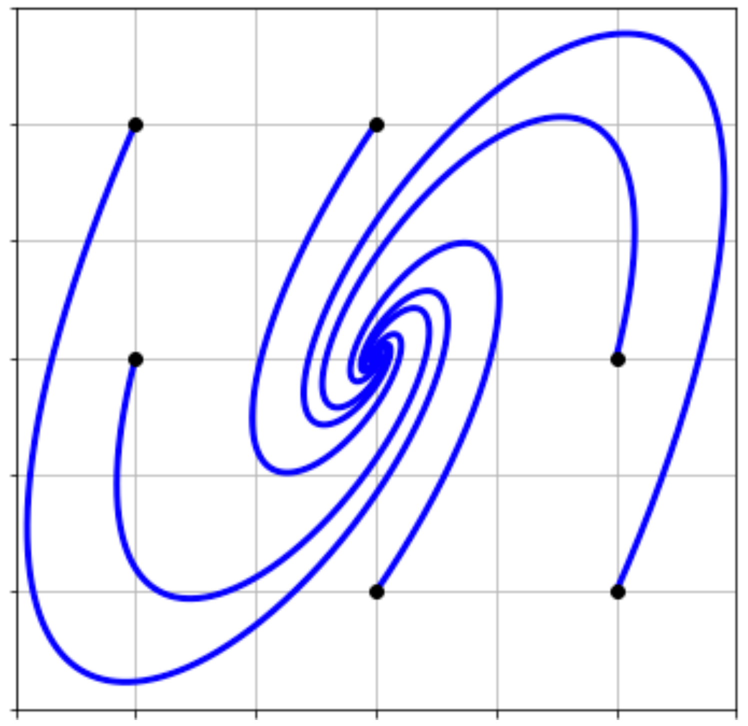
$$D\omega = \begin{bmatrix} - & \\ & + \end{bmatrix}, \quad \text{spec}(D\omega) \subset \mathbb{C}_+^\circ$$

Example:

$$f_1(x_1, x_2) = -x_1^2 + 4x_1x_2$$
$$f_2(x_1, x_2) = 6x_2^2 - 8x_1x_2$$

$$D\omega = \begin{bmatrix} -2 & 4 \\ -8 & 12 \end{bmatrix} \qquad \text{spec}(D\omega) = \{2 \pm 4i\}$$

# **Non-Nash stable equilibria:** saddle point




$$D\omega = \begin{bmatrix} - & \\ & + \end{bmatrix}, \quad \text{spec}(D\omega) \subset \mathbb{C}_+^\circ$$

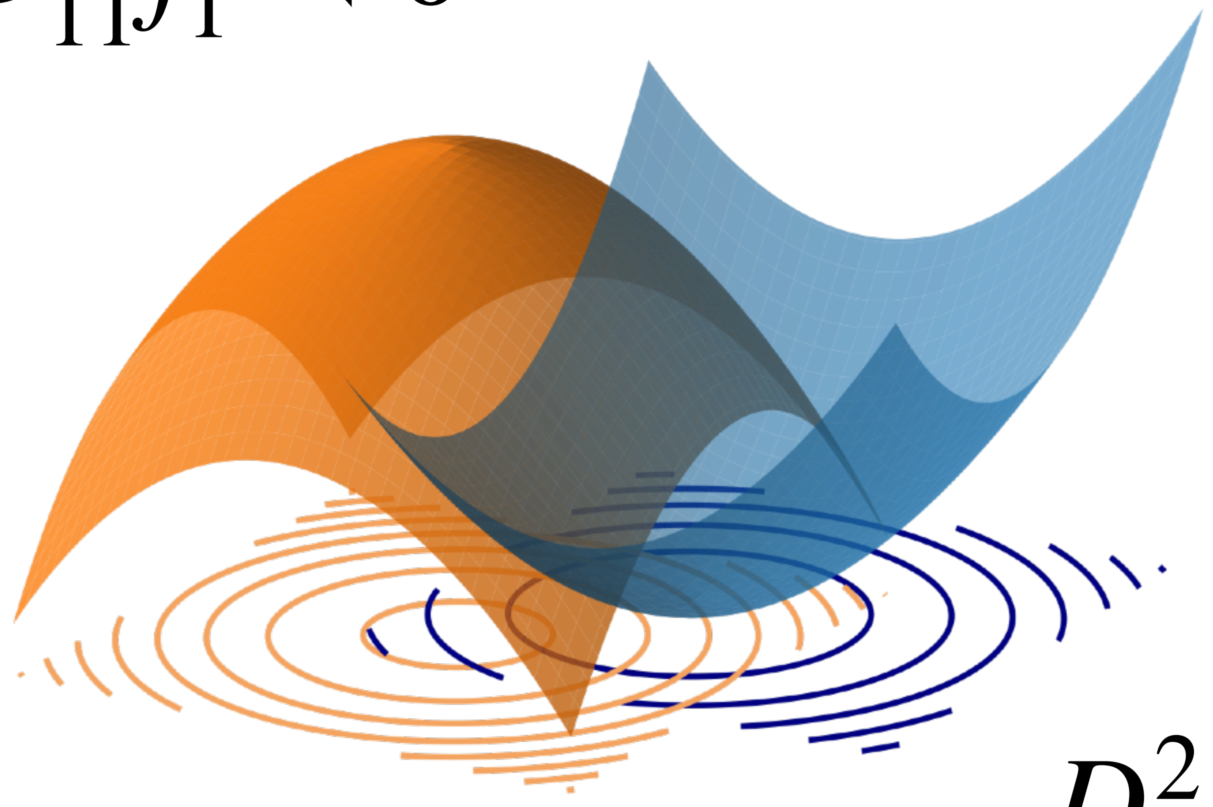
Example:

$$f_1(x_1, x_2) = -x_1^2 + 4x_1x_2$$
$$f_2(x_1, x_2) = 6x_2^2 - 8x_1x_2$$

$$D\omega = \begin{bmatrix} -2 & 4 \\ -8 & 12 \end{bmatrix}$$

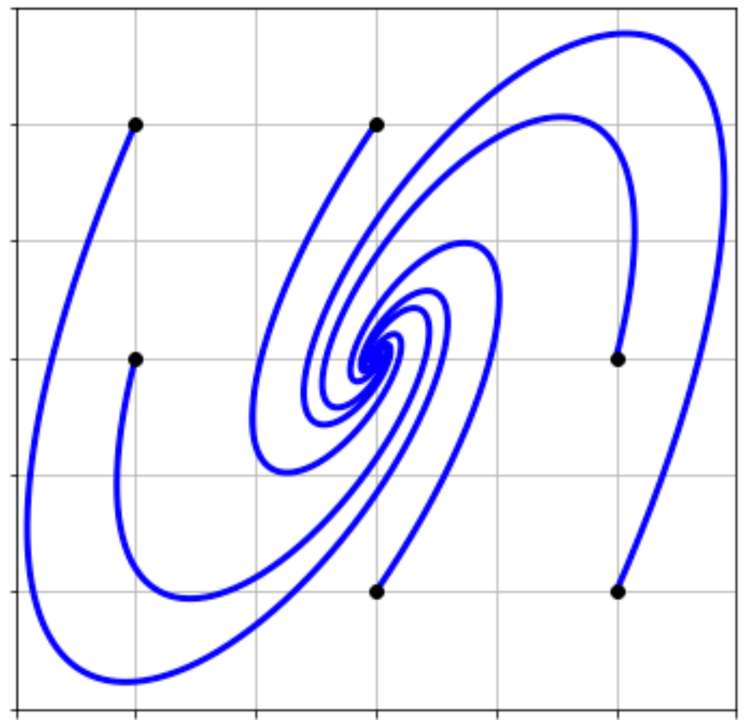
$$\text{spec}(D\omega) = \{2 \pm 4i\}$$

Agent 1 is at a maximum! $D_{11}^2 f_1 \prec 0$



$D_{22}^2 f_2 \succ 0$

# **Non-Nash stable equilibria:** saddle point



$\dot{x} = -\,\omega(x)$

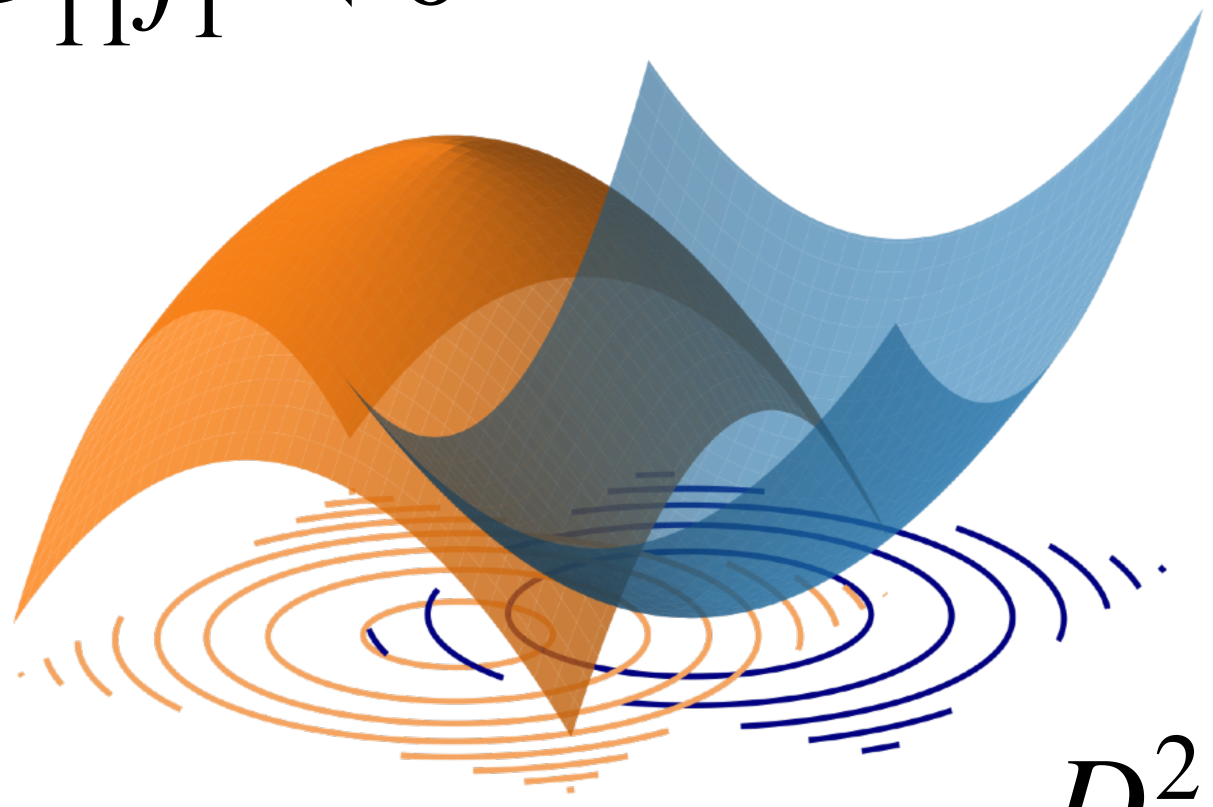
$$D\omega = \begin{bmatrix} - & \\ & + \end{bmatrix}, \quad \text{spec}(D\omega) \subset \mathbb{C}_+^{\circ}$$

Example:

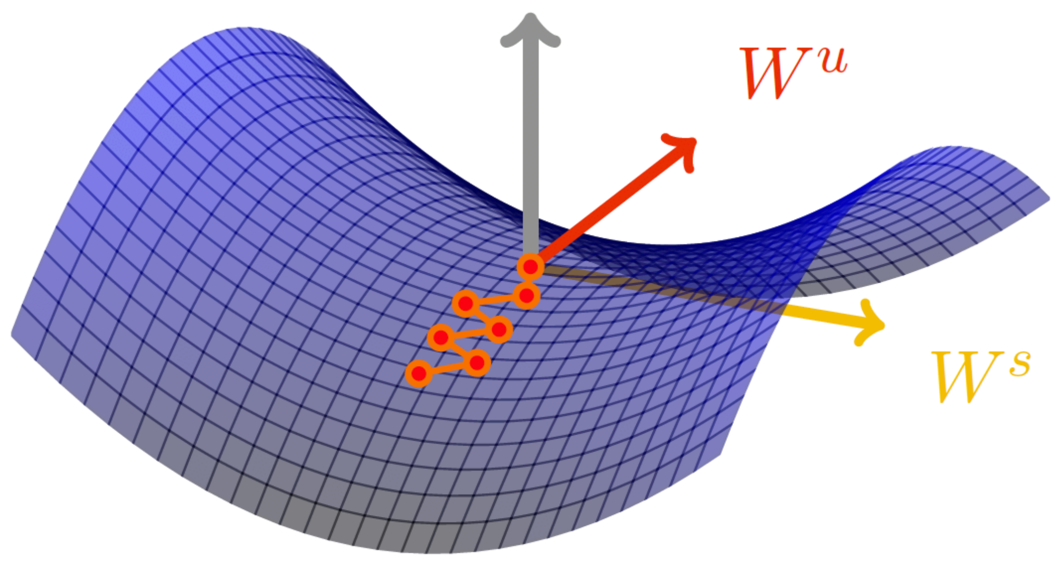
$$f_1(x_1, x_2) = -\,x_1^2 + 4x_1x_2$$
$$f_2(x_1, x_2) = 6x_2^2 - 8x_1x_2$$

$$D\omega = \begin{bmatrix} -2 & 4 \\ -8 & 12 \end{bmatrix} \qquad \text{spec}(D\omega) = \{2 \pm 4i\}$$

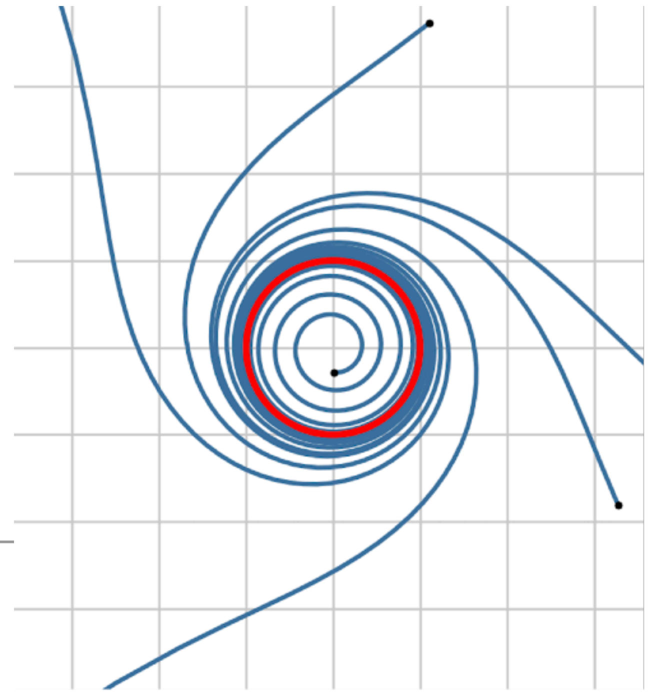
Agent 1 is at a maximum! $D_{11}^2 f_1 \prec 0$



$D_{22}^2 f_2 \succ 0$



# **Periodic orbit:** LQ game policy gradient



Each player aims to minimize their own cost

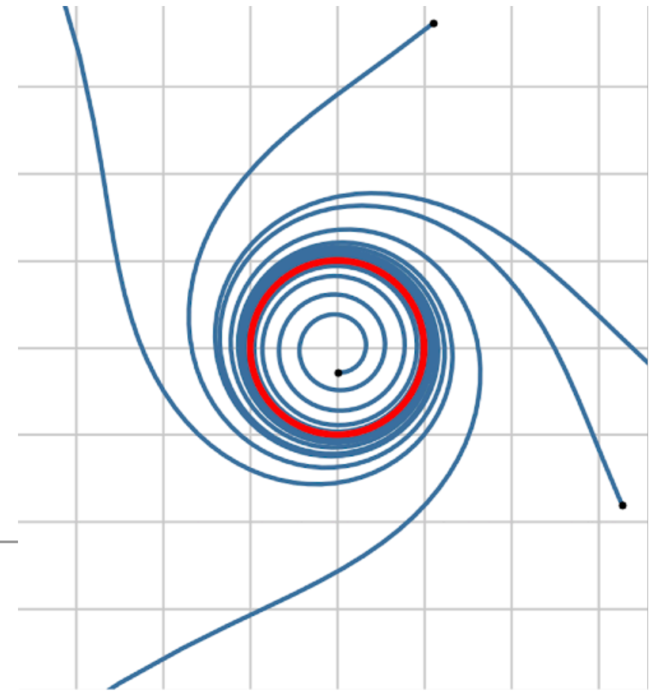
$$J_i(u_1, u_2) = \sum_t x_t^T Q_i z_t + u_t^T R_i u_t, \ \ i = 1, 2 \ \ \ u_t = [u_{1,t} \ u_{2,t}]^T$$

subject to $z_{t+1} = Az_t + B_1 u_{1,t} + B_2 u_{2,t}$

- Parameterize feedback controllers: $u_i = -K_i z_t$
- Nash policies $(K_1, K_2)$ exist and are unique
- Policy gradient:

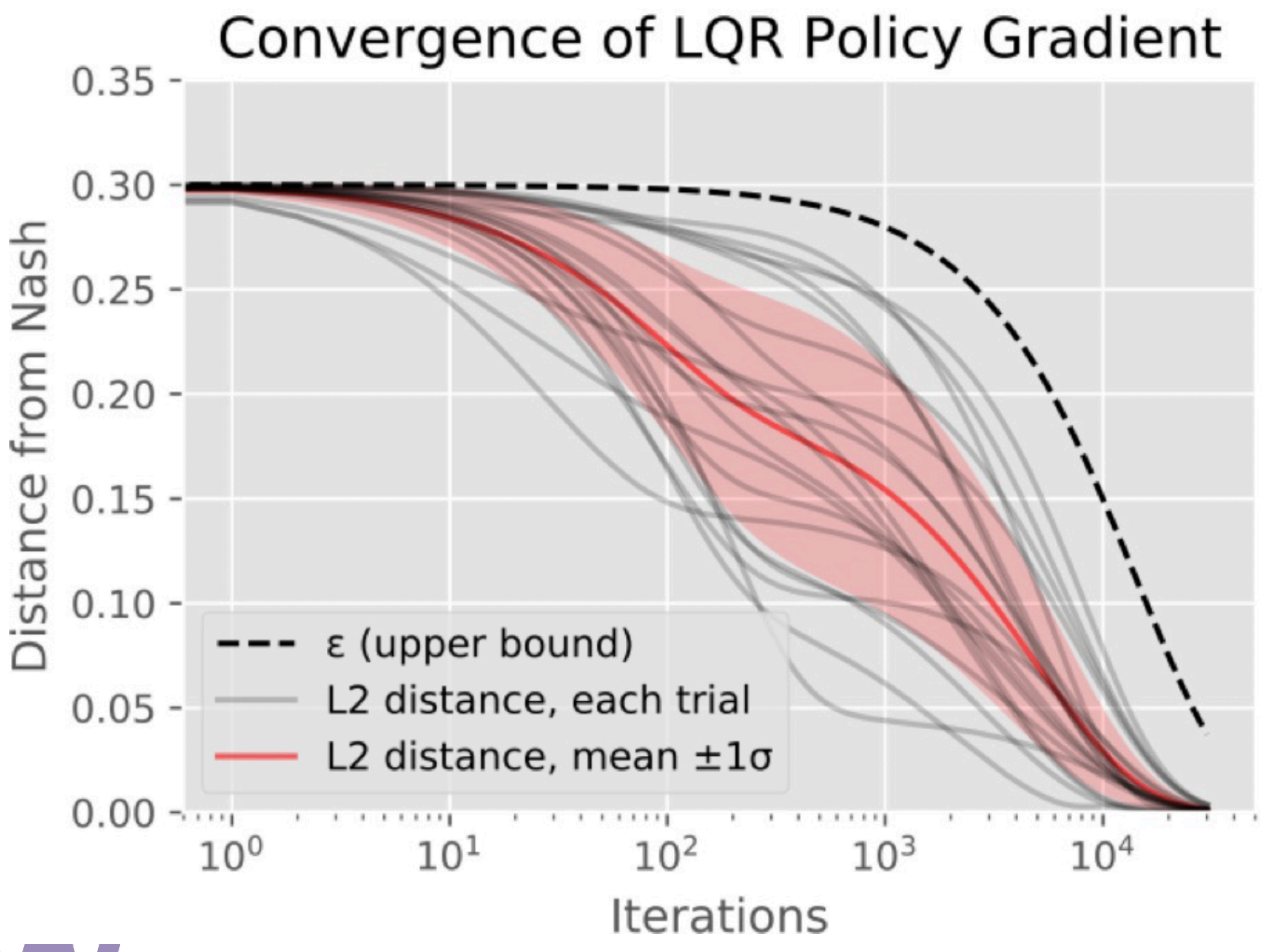
$$K_i^+ = K_i - \eta_i \widehat{D_i J_i}(K_1, K_2)$$

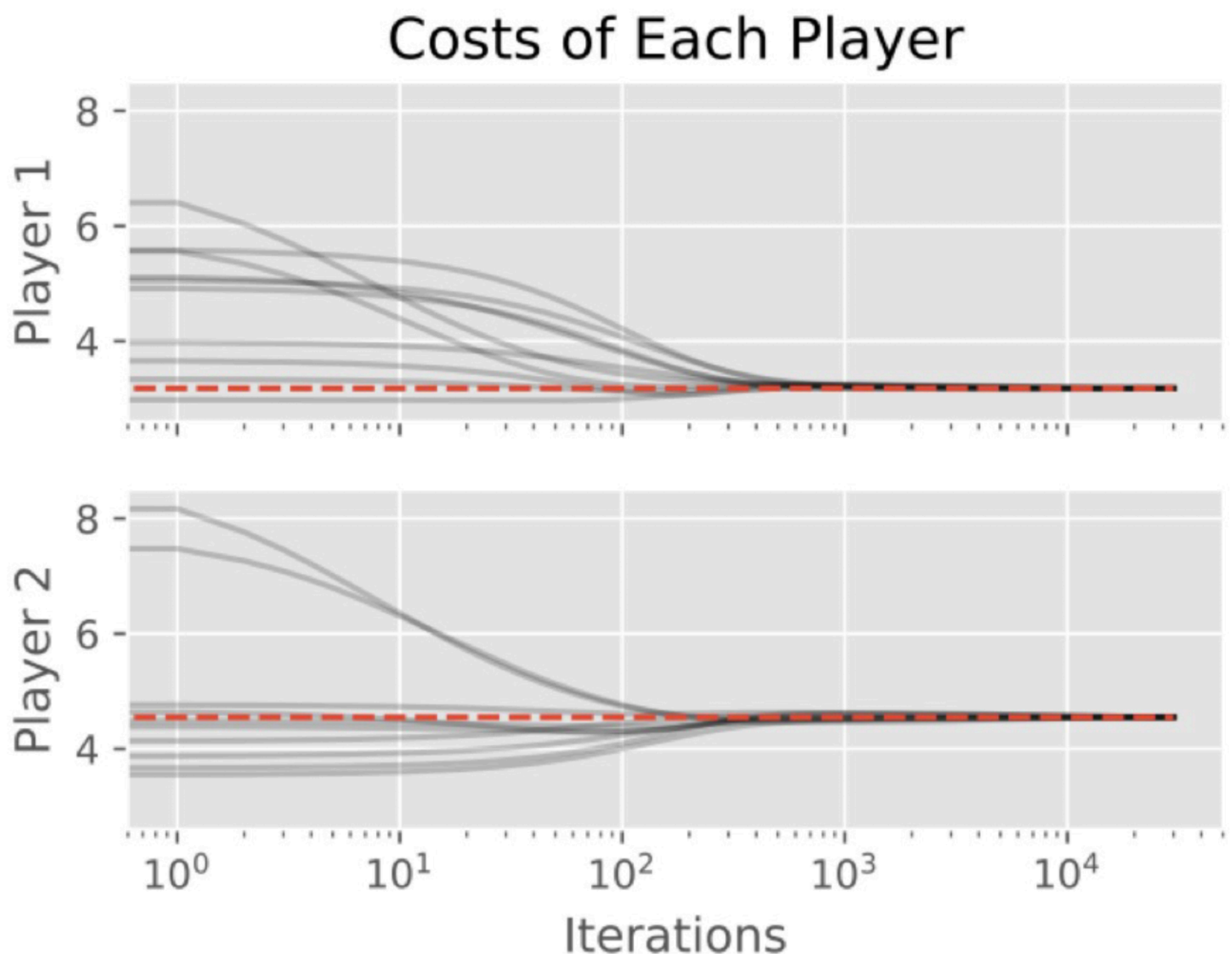
# **Periodic orbit:** LQ game policy gradient

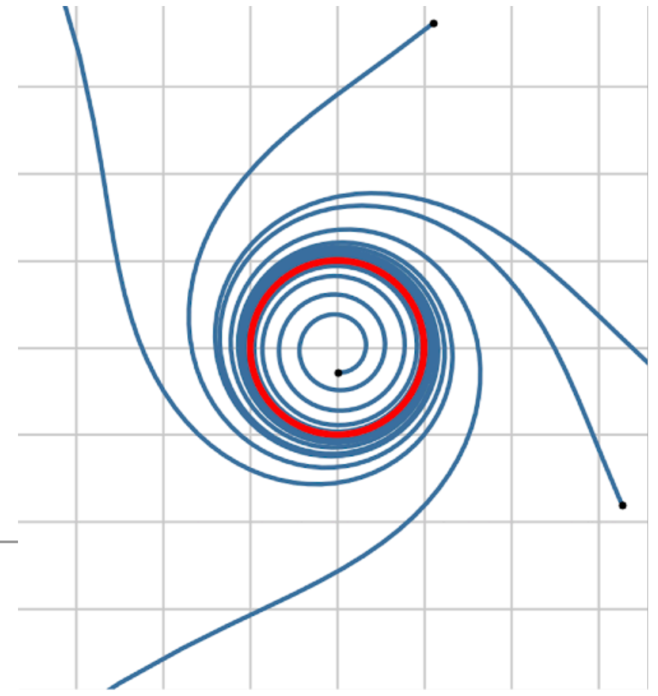


- Stable parameters

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0.4 \end{bmatrix},\ B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ Q_{1,2} = I,\ R_1 = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix},\ R_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$



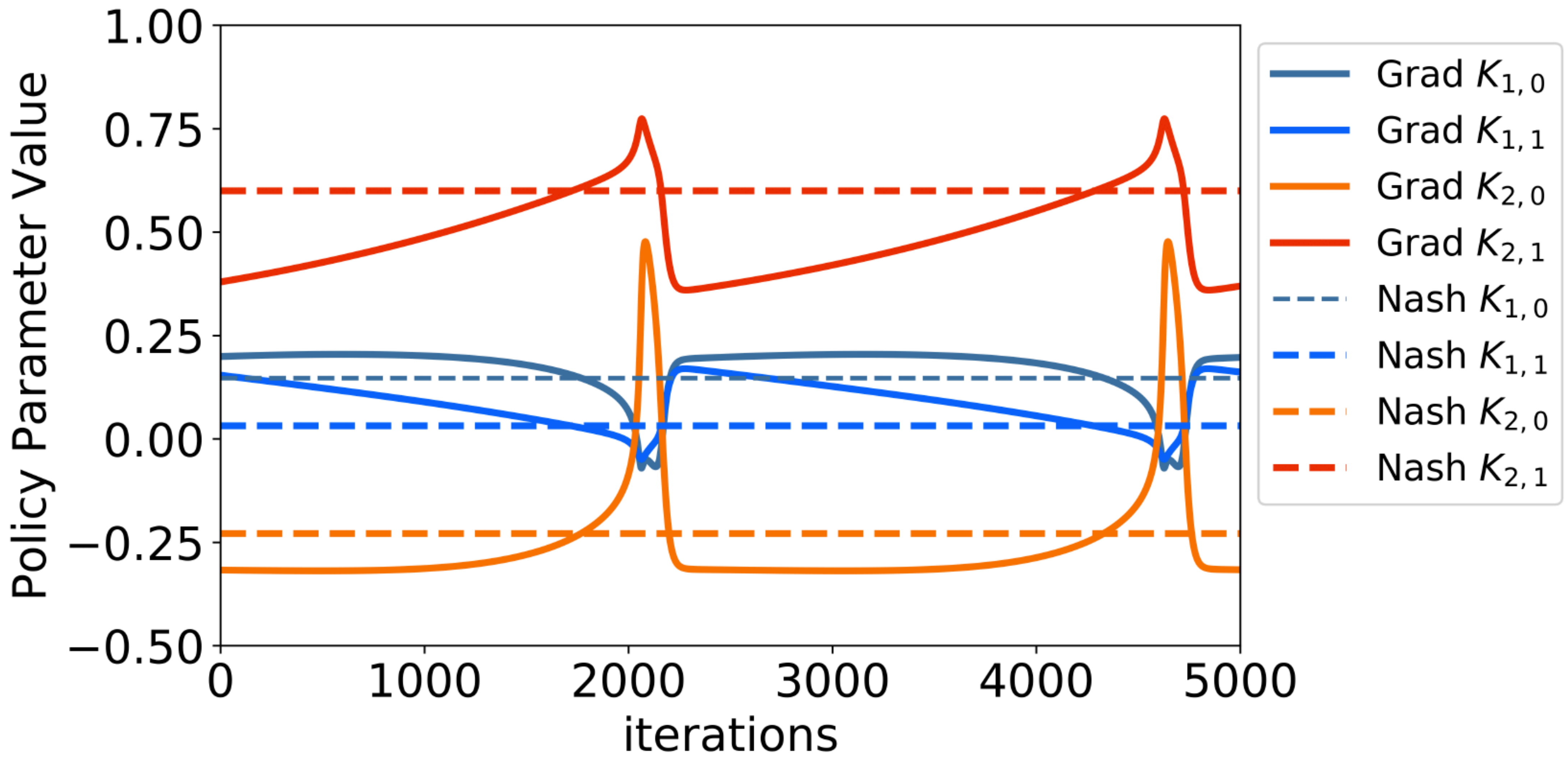


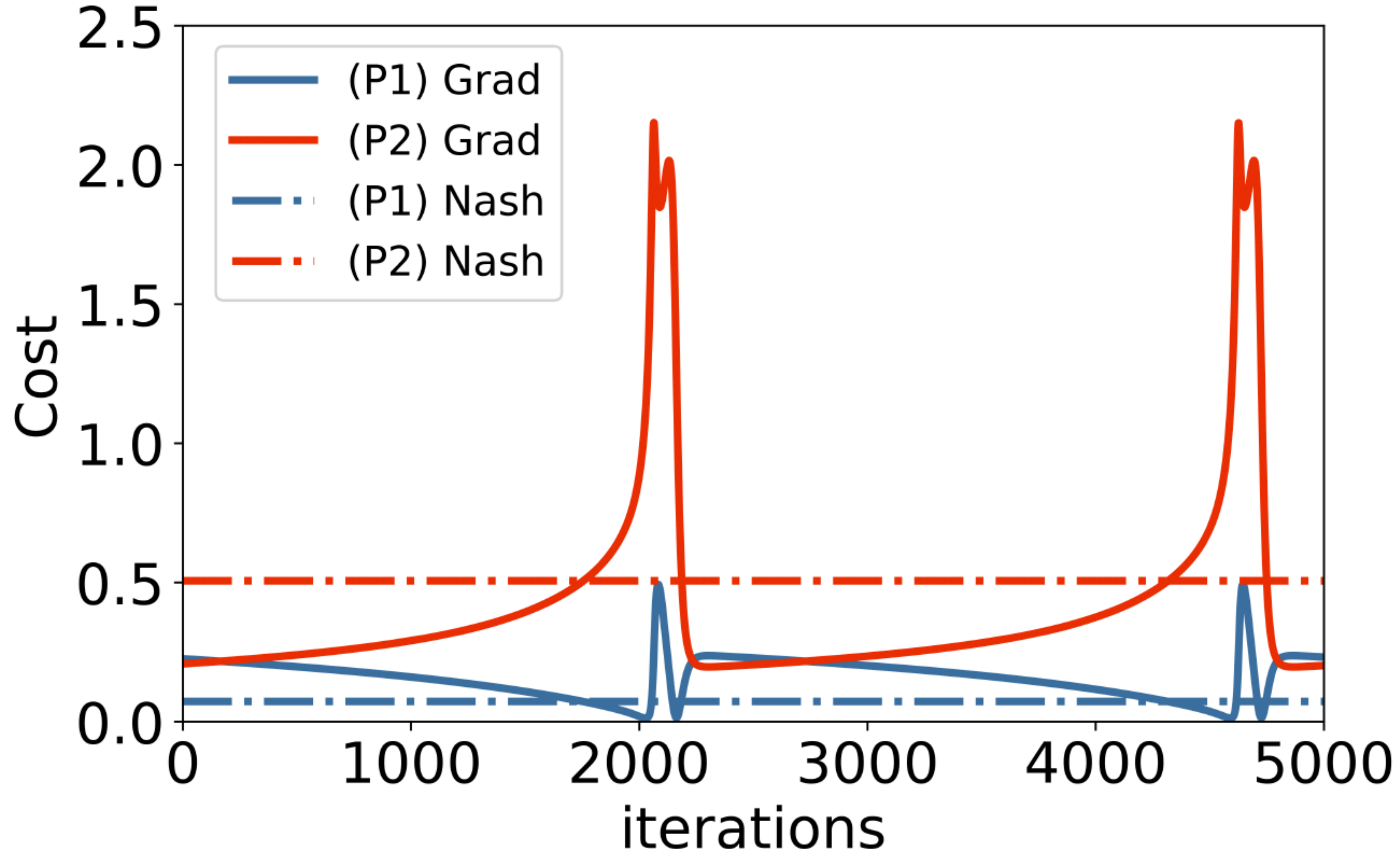
# **Periodic orbit:** LQ game policy gradient

- Unstable parameters!

$$z_{t+1} = \underbrace{\begin{bmatrix} 0.75 & 0.001 \\ 0.02 & 0.66 \end{bmatrix}}_{A} z_t + \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{B_1} u_1 + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{B_2} u_2$$

cost parameters: $Q_1 = \begin{bmatrix} 0.01 & 0 \\ 0 & 1 \end{bmatrix},\ Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0.15 \end{bmatrix},\ R_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ R_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix}$



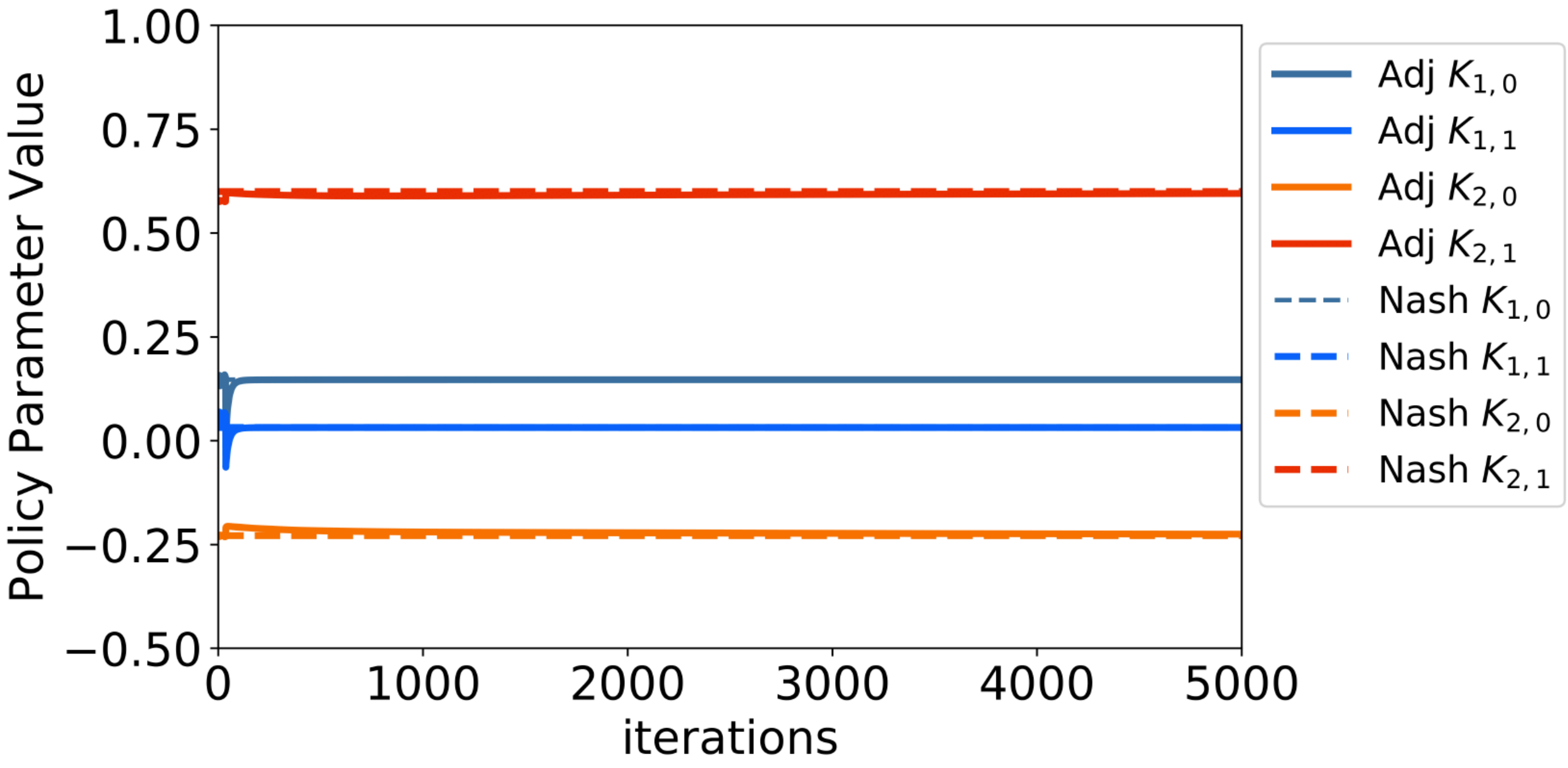



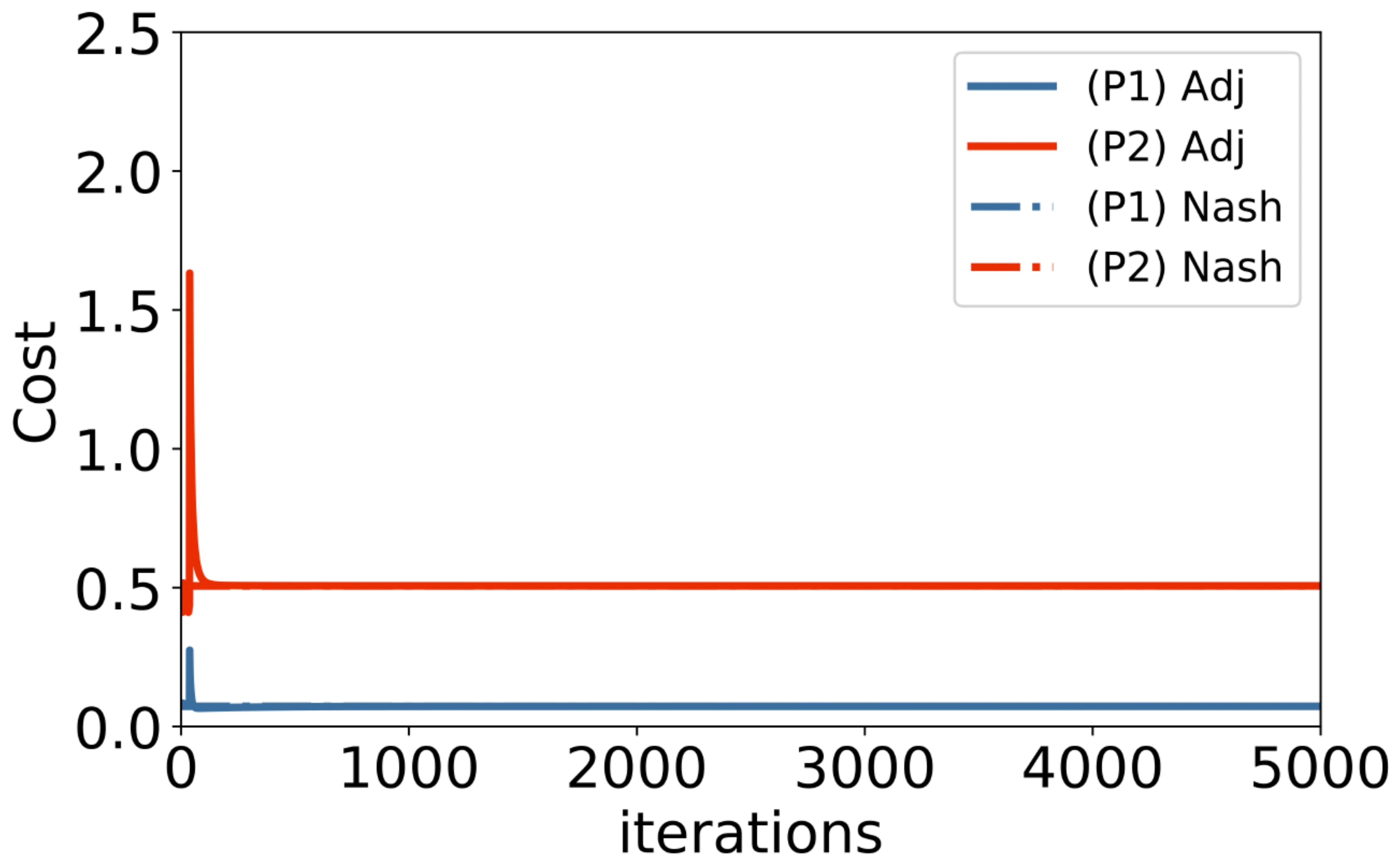
# Solution: modified learning rule via coordination

$$\dot{x} = -\frac{1}{2}\left(\omega(x) + \tilde{H}(x)D\omega^{-1}(x)\omega(x)\right)$$

where

$$\tilde{H} = \text{diag}(D_{11}^2 f_1,\ \cdots,\ D_{nn}^2 f_n) - D\omega$$



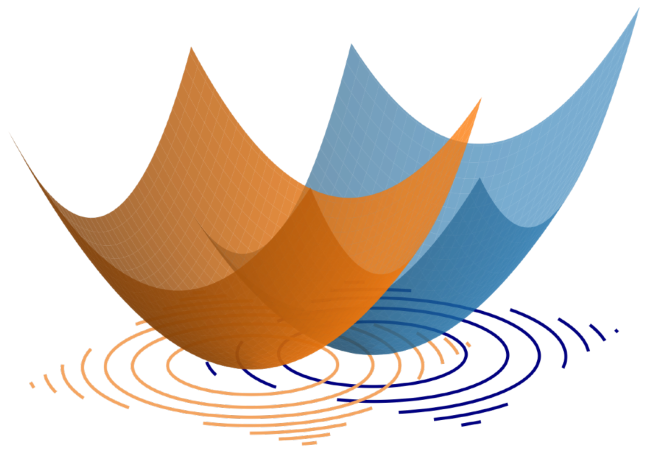




Mazumdar, Sastry, Jordan. On Finding Local Nash equilibria (and Only Local Nash equilibria) in Zero-Sum Games, arxiv:1901.00838.

# Conclusion

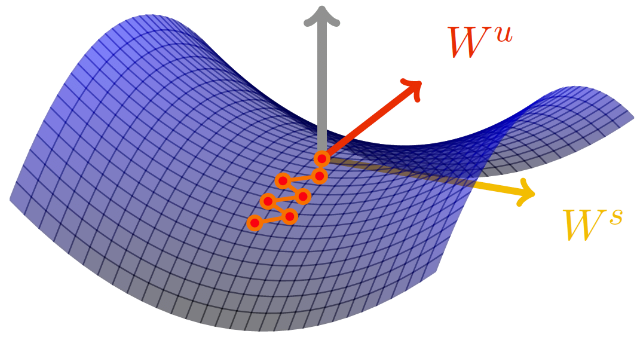
- **Local convergence**

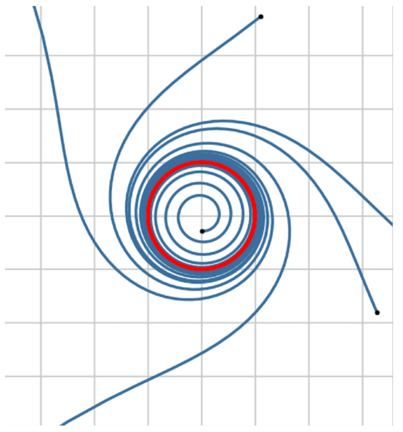
✅ Nash equilibria



- **“Undesirable” limiting behaviors**

⭕ Non-Nash ⭕ Periodic

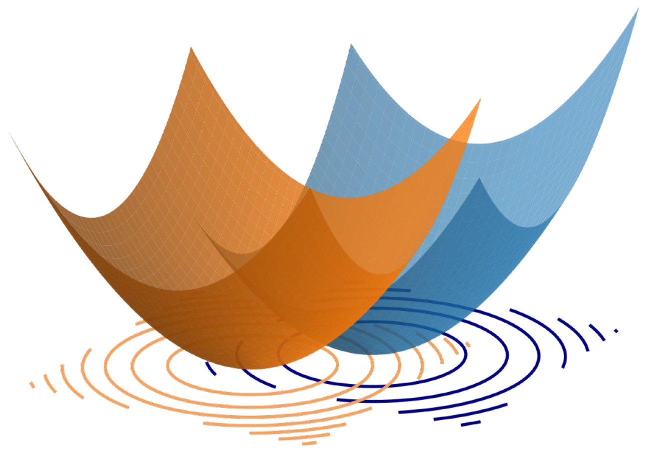


# Conclusion

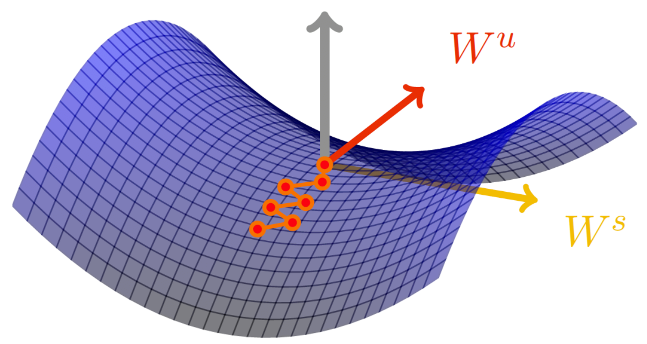
- **Local convergence**

✅ Nash equilibria

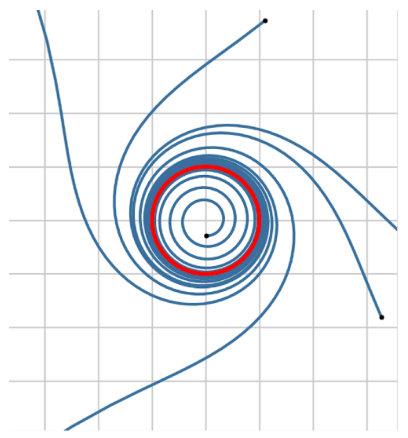


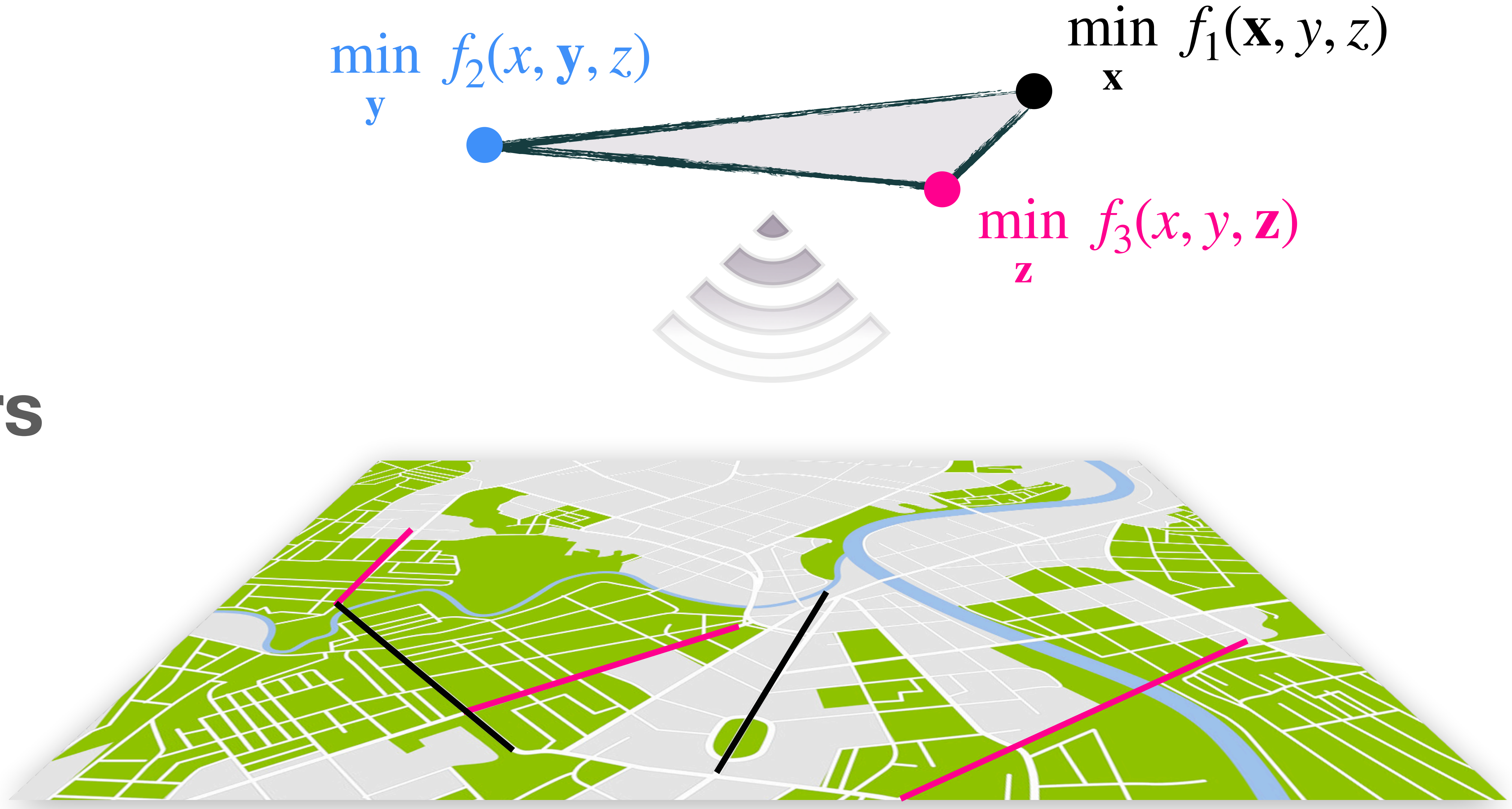
- **"Undesirable" limiting behaviors**

⭕ Non-Nash ⭕ Periodic





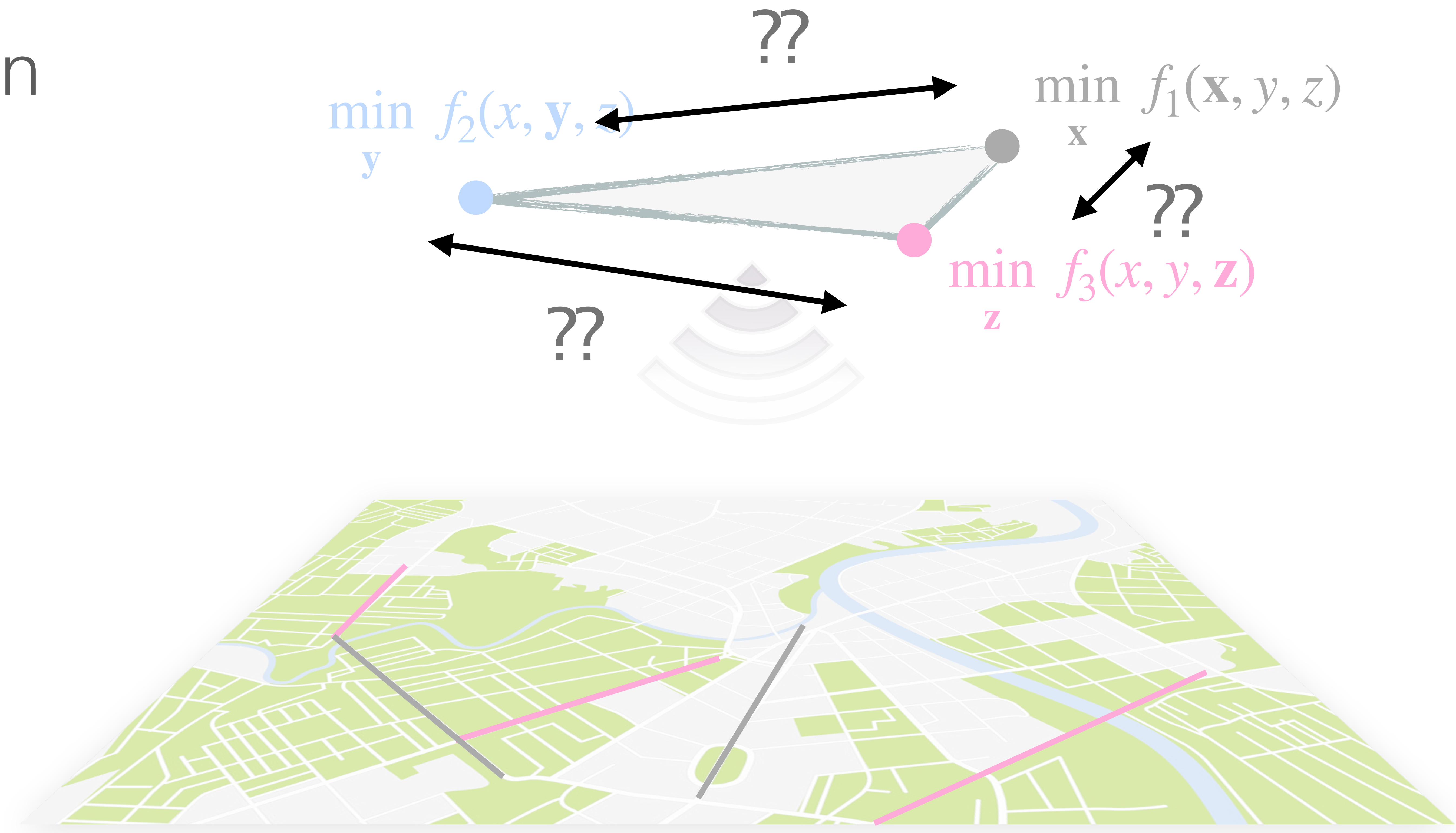




# Future work

- **Efficient coordination methods**

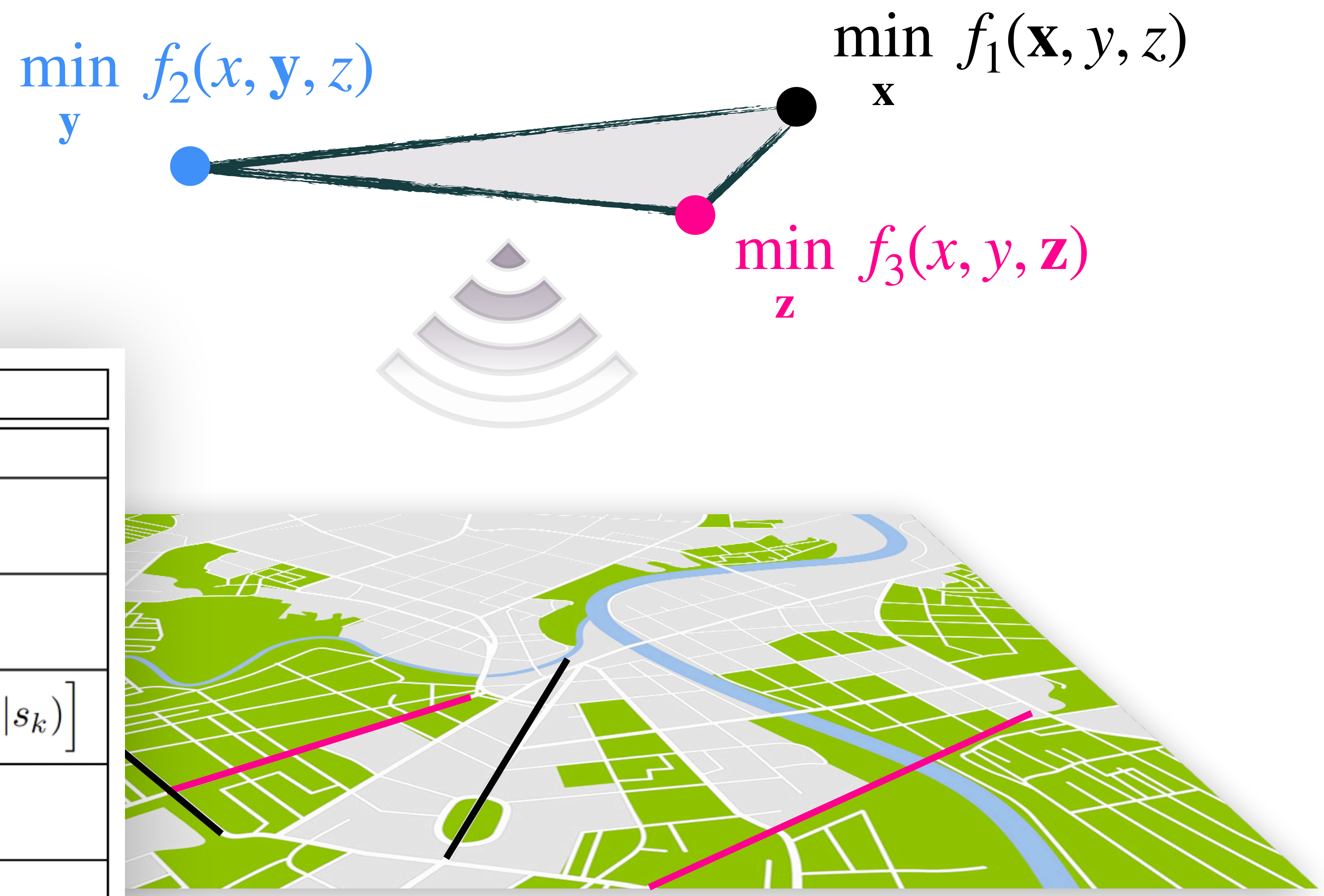
  Sharing other gradient information $D_i f_j^{(i \neq j)}$ to converge to Nash?



# Future work

- Efficient coordination methods
- **Other updates rules (stochastic):**
  Do the dynamics correspond to the original game?

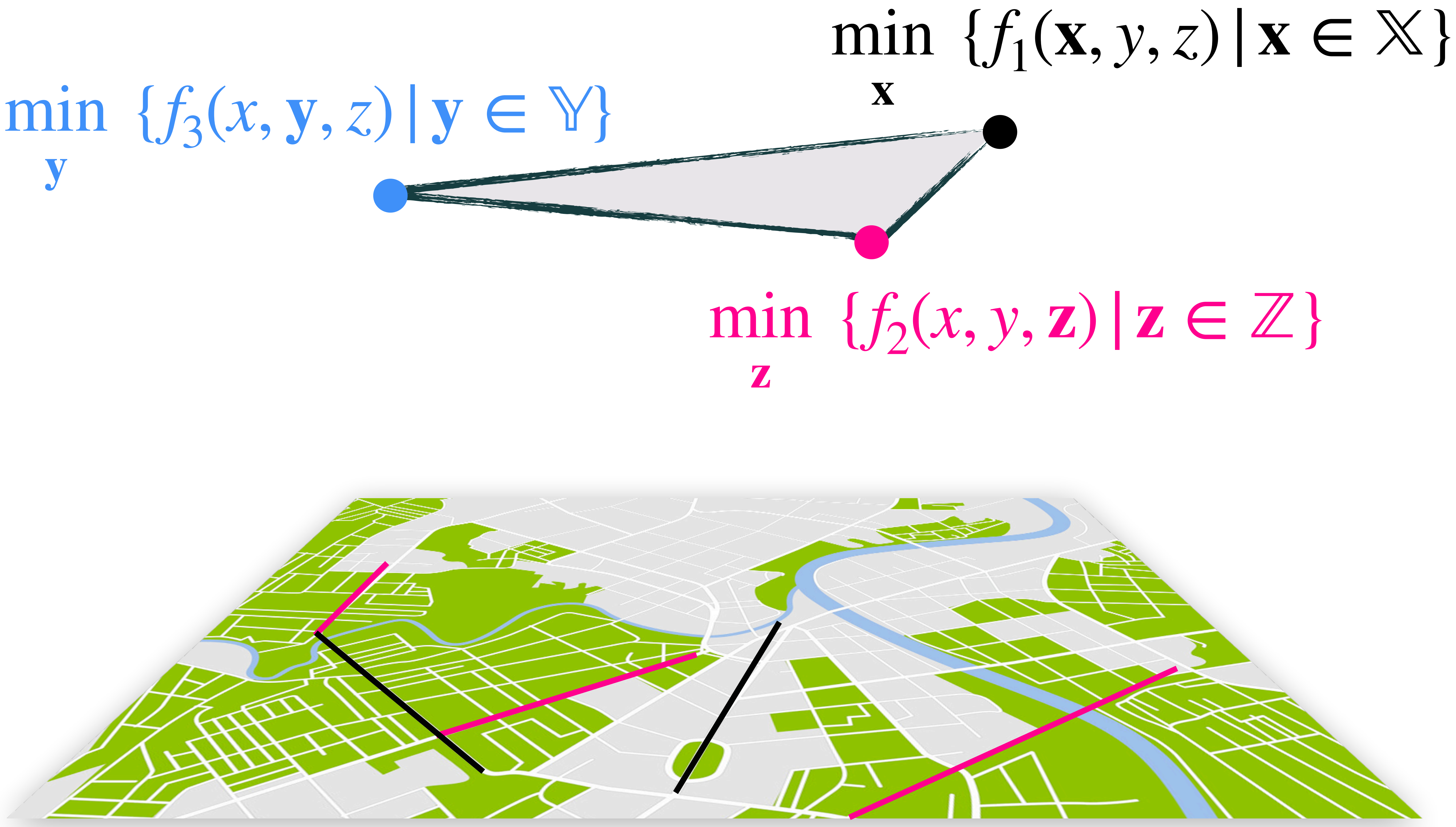
| | Problem Class | Gradient Learning Rule |
|---|---|---|
| | Gradient Play | $x_i^+ = x_i - \gamma_i D_i f_i(x_i, x_{-i})$ |
| ZS | GANs | $\theta^+ = \theta - \gamma \nabla_\theta \ell(\theta, w)$<br>$w^+ = w + \gamma \nabla_w \ell(\theta, w)$ |
| MARL | Policy Gradient w/ Policy Prediction | $x_1^+ = x_1 - \gamma_1 \nabla_1 J_1(\pi_1(x_1), \pi_2(x_2) + \delta \nabla_2 J_2(\pi_1(x_1), \pi_2(x_2)))$ |
| | Multi-Agent Policy Gradient | $x_i^+ = x_i - \gamma_i \mathbb{E}_{\tau \sim P_T(\pi)} \left[ \sum_t R_i(s_t, u_t) \sum_{k=0}^{t} \nabla_i \log \pi_i(x_i)(u_{i,k} \vert s_k) \right]$ |
| MAB | Individual Q-Learning | $q_i^+(u_i) = q_i(u_i) + \gamma_i(r_i(u_i, \pi_{-i}(q_i, q_{-i})) - q_i(u_i))$ |
| | Multi-Agent Gradient Bandits | $x_{i,\ell}^+ = x_{i,\ell} + \gamma_i \mathbb{E}[\beta_i R_i(u_i, u_{-i}) \vert u_i = u], \; u \in U_i$ |
| | Multi-Agent Experts | $x_{i,\ell}^+ = x_{i,\ell} + \gamma_i \mathbb{E}[R_i(u_i, u_{-i}) \vert u_i = u], \; u \in U$ |



# Future work

- Efficient coordination methods
- Other update rules (stochastic)
- **Constrained optimization**

    Safety constraints



# Questions?

- Thank you!

$$y^+ = y - \gamma D_2 f_2(x, \mathbf{y}, z)$$

$$x^+ = x - \gamma D_1 f_1(\mathbf{x}, y, z)$$

$$z^+ = z - \gamma D_3 f_3(x, y, \mathbf{z})$$

UNIVERSITY of WASHINGTON

COMPUTATIONAL NEUROSCIENCE CENTER

NSF

**Theorem**: ($x^*$: stable differential Nash)
suppose $x_0 \in B_r(x^*)$, $\omega$ is Lipschitz, and $\gamma_i = \sqrt{\alpha}/(k\beta)$ for each $i \in [n]$ with $\alpha < k\beta$. Gradient based learning obtains an $\varepsilon$-differential Nash in finite time $T \geq \lceil 2k\frac{\beta}{\alpha}\log(r/\varepsilon)\rceil$

$$\alpha = \min_{x\in B_r(x)} \sigma^2_{\min}(\underbrace{D\omega(x) + D\omega(x)^T}_{\text{symmetric part of } D\omega}),$$

$$\beta = \max_{x\in B_r(x)} \sigma^2_{\max}(D\omega(x))$$