Opinion Dynamics for Decision Making and Learning in Multi-agent Interactions

Naomi Ehrich Leonard
Mechanical and Aerospace Engineering
Princeton University
naomi@princeton.edu
Papers: https://naomi.princeton.edu/publications/

Joint work with

Anastasia Bizyaeva
Mechanical and Aerospace Engineering
Princeton University

Alessio Franci
Department of Mathematics
National Autonomous University of Mexico

Shinkyu Park
Electrical and Computer Engineering
KAUST

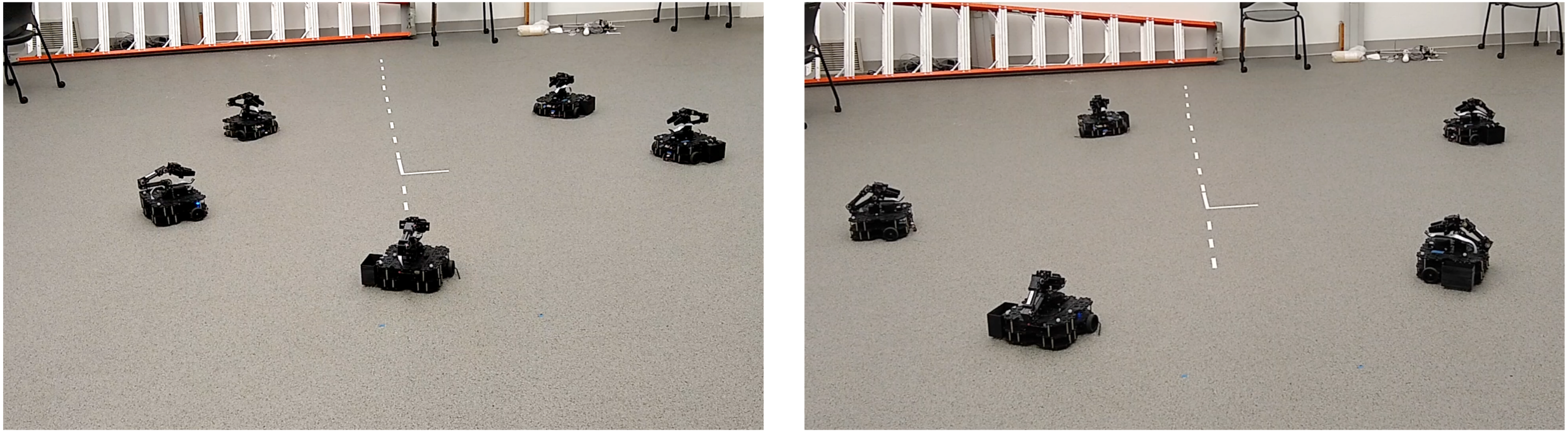
Big Questions for Network of Distributed Decision-Making Agents

1. How to ensure quick, reliable, and informed decision-making in response to external cues?

2. How to enable sufficiently rich suite of behaviors to meet demands of mission and environment?

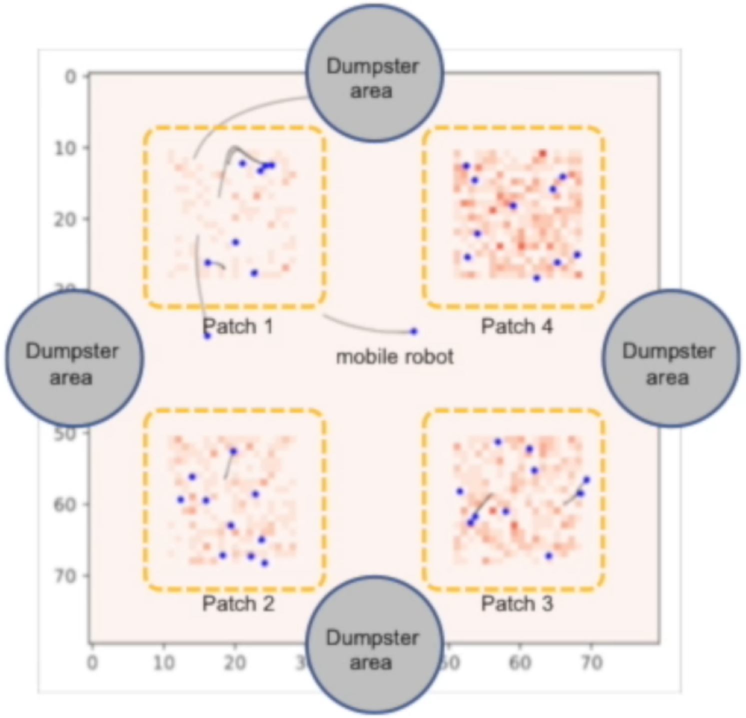
3. What role does network structure play in transient and steady state? How to leverage in design?

Motivating Design Problem: Cooperative Navigation

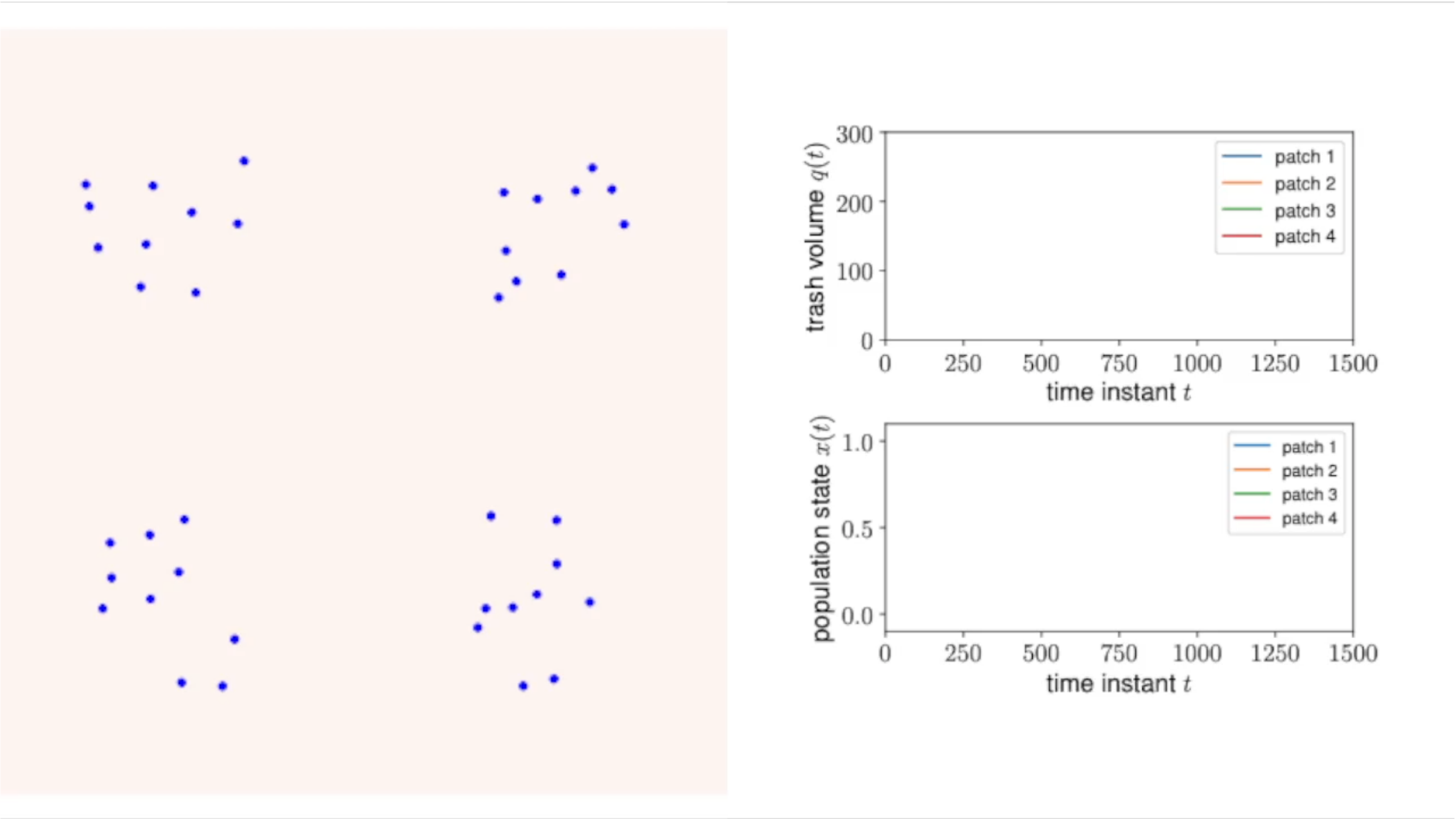


Videos: Shinkyu Park

Motivating Design Problem: Dynamic Task Allocation



Multi-Robot Trash Collection Environment



Video: Shinkyu Park

Park, Zhong, and Leonard, “Multi-robot task allocation games in dynamically changing environments,” ICRA, 2021

Popular model for consensus: Weighted average update (DeGroot)

$x_i \in \mathbb{R}$ is opinion of agent i for $i = 1, \dots, N_a$

$$a_{i1} + \cdots + a_{iN_{\mathrm{a}}} = 1$$

$$x_i(t+1) = a_{i1}x_1(t) + \cdots + a_{iN_{\mathrm{a}}}x_{N_{\mathrm{a}}}(t)$$

Equivalently,

$$x_i(t+1) = x_i(t) + \Big(- x_i(t) + a_{i1}x_1(t) + \cdots + a_{iN_{\mathrm{a}}}x_{N_{\mathrm{a}}}(t) \Big)$$

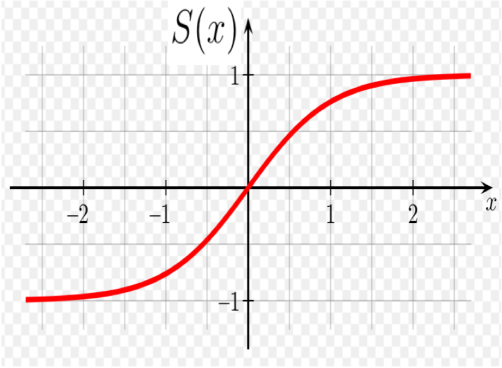
Discretization of continuous linear consensus dynamics

$$\dot{x}_i = \underset{\substack{\nearrow \\ \text{damping} \\ \text{(neg feedback)}}}{-x_i} + \underbrace{a_{i1}x_1 + \cdots + a_{iN_{\mathrm{a}}}x_{N_{\mathrm{a}}}}_{\substack{\text{opinion exchange} \\ \text{(pos feedback)}}}$$

Nonlinear model for opinion formation

Opinion exchanges are saturated:

$$\dot{x}_i = -x_i + S(a_{i1}x_1 + \cdots + a_{iN_\mathrm{a}}x_{N_\mathrm{a}})$$



Attention parameter $u_i \geq 0$ and additive input b_i are introduced:

$$\dot{x}_i = -\underset{\substack{\uparrow\\ \text{damping}\\ \text{(neg feedback)}}}{d_i} x_i + u_i \underbrace{S(a_{i1}x_1 + \cdots + a_{iN_\mathrm{a}}x_{N_\mathrm{a}})}_{\substack{\text{opinion exchange}\\ \text{(pos feedback)}}} + b_i$$

For **small** $u_i \implies$ system behaves **linearly**: $x_i \approx b_i$

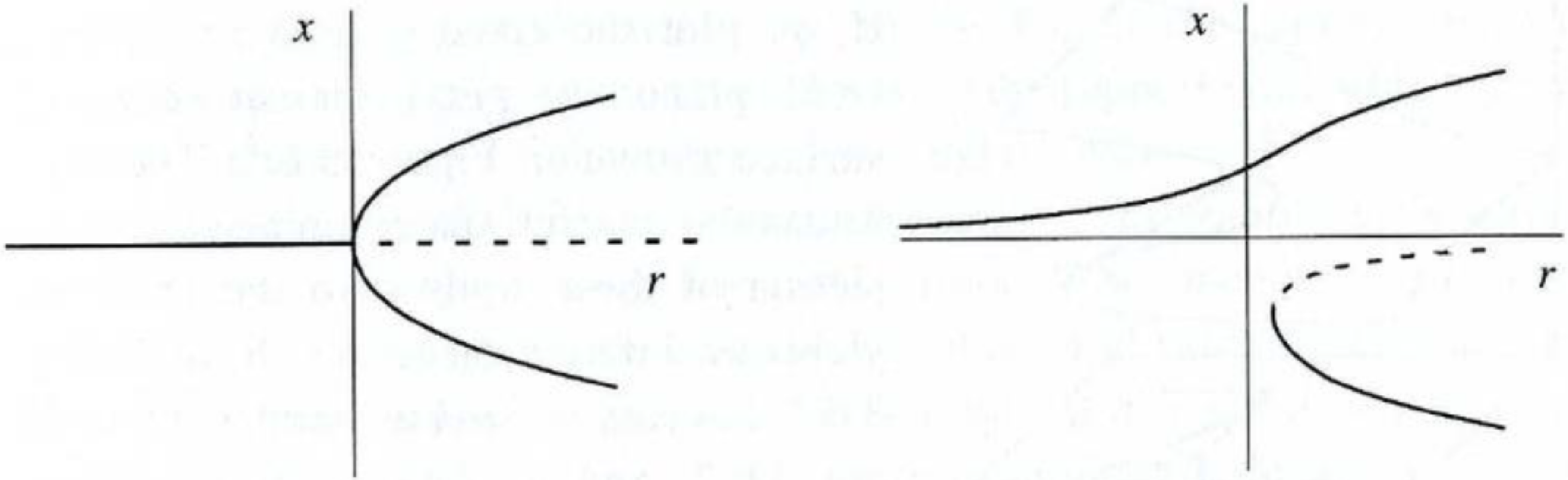
For **large** $u_i \implies$ system behaves **nonlinearly**: $|x_i| \gg |b_i|$

Bizyaeva, Franci, Leonard, “A general model of opinion dynamics with tunable sensitivity”, arXiv:2009.04332, Oct 2020

Bifurcation

$$\dot{x} = f(x, r)$$

A **bifurcation** is a qualitative change in the number, configuration and/or stability of equilibria of a system as a parameter is varied



Nonlinear multi-option opinion formation

Set of N_a agents: $\{1, \ldots, N_a\}$ and set of N_o options: $\{1, \ldots, N_o\}$

Each agent i has opinion of each option j: $\hat{z}_{ij} \in \mathbb{R}$

Higher opinion corresponds to greater $\hat{z}_{ij}$

Each agent i has *relative* opinion of each option j: $z_{ij} = \hat{z}_{ij} - \sum_{l \neq j}^{N_o} \hat{z}_{il} \implies z_{i1} + \cdots z_{iN_o} = 0$

$$\dot{z}_{ij} = F_{ij}(\boldsymbol{Z}) - \frac{1}{N_\mathrm{o}} \sum_{l=1}^{N_\mathrm{o}} F_{il}(\boldsymbol{Z})$$

$$F_{ij}(\boldsymbol{Z}) = \underbrace{-d_{ij} z_{ij}}_{\substack{\text{damping} \\ \text{(neg feedback)}}} + u_i \underbrace{\sum_{l=1}^{N_o} S_{ijl} \left(\sum_{k=1}^{N_\mathrm{a}} A_{ik}^{jl} z_{kl} \right)}_{\substack{\text{opinion exchange} \\ \text{(pos feedback)}}} + b_{ij}$$

Bizyaeva, Franci, Leonard, 2020

Two-option opinion formation

Interpret consensus dynamics and nonlinear dynamics (with saturation) as opinion dynamics on 2 options

$$\boldsymbol{Z}_i = (z_{i1}, z_{i2}), \text{ with } z_{i1} + z_{i2} = 0$$

$$x_i = z_{i1} \text{ and } -x_i = z_{i2}$$

$$\Longrightarrow$$

$x_i > 0$, agent i prefers option 1

$x_i = 0$, agent i is neutral

$x_i < 0$, agent i prefers option 2

Nonlinear two-option opinion formation with homogeneity

$$\dot{x}_i = -dx_i + u_i S\left(\alpha x_i + \gamma \sum_{\substack{k=1\\k\neq i}}^{N_{\mathrm{a}}} a_{ik} x_k\right) + b_i$$

$\alpha \geq 0$ is self-reinforcing weight

$\gamma \in \mathbb{R}$ is inter-agent weight

$b_i = \frac{1}{2}(b_{i1} - b_{i2}) \in \mathbb{R}$ is input (evidence for option 1 if $b_i > 0$)

$A = [a_{ik}]$ is unweighted adjacency matrix of network graph

Agents i and k are *cooperative* if $\gamma > 0$ and *competitive* if $\gamma < 0$

Bizyaeva, Franci, Leonard, 2020

Definitions

λ_{max} is eigenvalue of A with largest real part; $\mathbf{v}_{max}$ corresponding unit left eigenvector

λ_{min} is eigenvalue of A with smallest real part; $\mathbf{v}_{min}$ corresponding unit left eigenvector

$W(\lambda_i)$ the generalized eigenspace associated to λ_i

Agreement equilibria: $x_i \neq 0$, $\text{sign}(x_i) = \text{sign}(x_k)$ for all i, k

Disagreement equilibria: $\text{sign}(x_i) = -\,\text{sign}(x_k)$ for at least one pair i, k, $i \neq k$

Special Case of Theorem 1 (Bizyaeva et al, 2020): Opinion Formation as Bifurcation

Let G be a connected undirected. The following hold with $u_i := u \geq 0$ and $b_i = 0$ for all $i = 1, \ldots, N_a$:

A. Cooperation leads to agreement:

If $\gamma > 0$, the neutral state $\mathbf{x} = 0$ is a locally exponentially stable equilibrium for $0 < u < u_a$ and unstable for $u > u_a$,

$$u_a = \frac{d}{\alpha + \gamma \lambda_{max}}.$$

At $u = u_a$, branches of agreement equilibria emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{max})$;

B. Competition leads to disagreement:

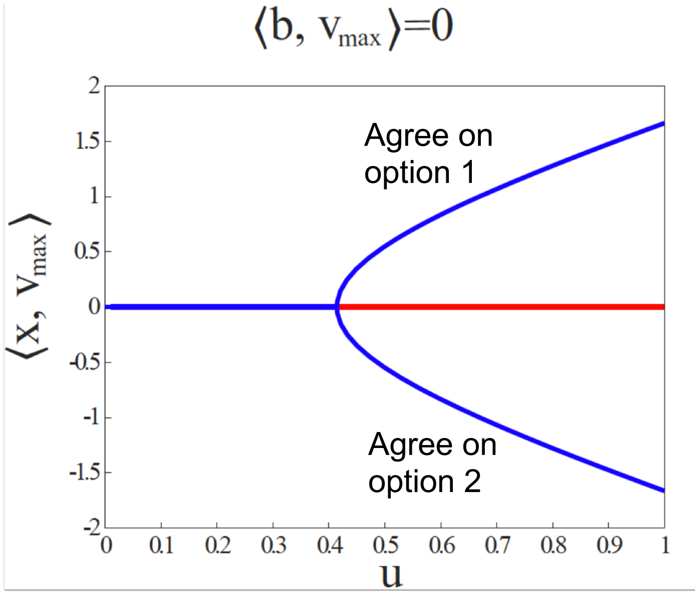
If $\gamma < 0$ the neutral state $\mathbf{x} = \mathbf{0}$ is a locally exponentially stable equilibrium for $0 < u < u_d$ and unstable for $u > u_d$,

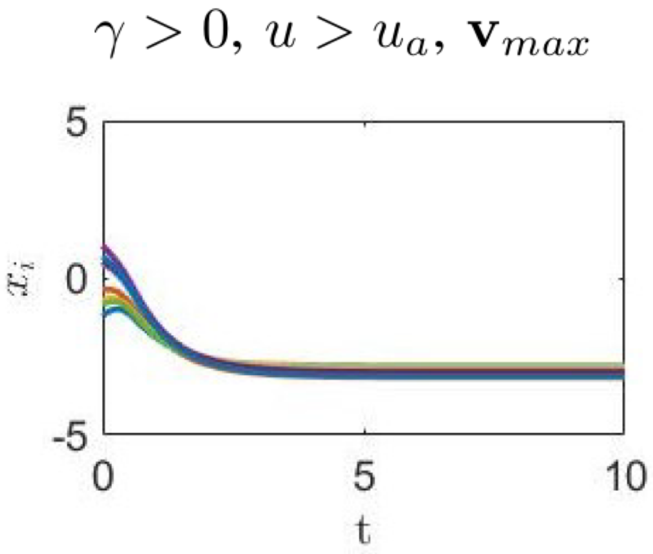
$$u_d = \frac{d}{\alpha + \gamma \lambda_{min}}.$$

At $u = u_d$, branches of disagreement equilibria emerge in a steady-state bifurcation off of $\mathbf{x} = \mathbf{0}$ along $W(\lambda_{min})$.

Theorem 2
Opinion Formation as Pitchfork Bifurcation

The agreement and disagreement bifurcations in Theorem 1 are supercritical pitchfork bifurcations.





Example:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\lambda_{max} = 1$$

$$\mathbf{v}_{max} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

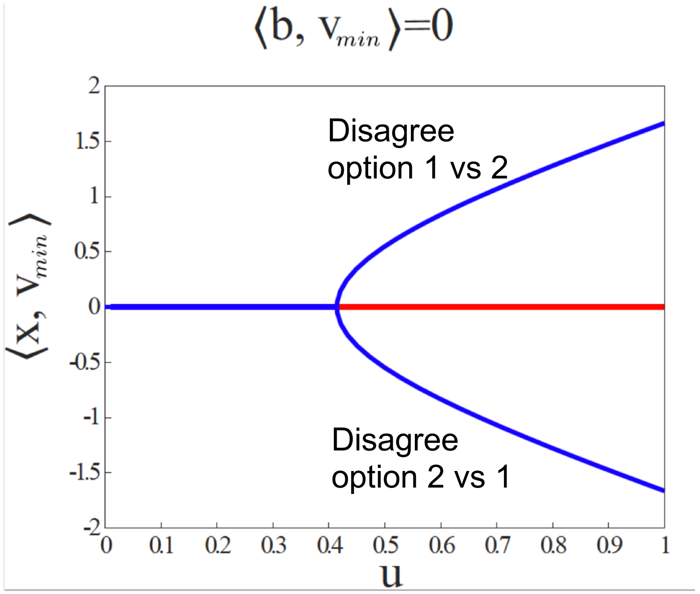
$$\langle \mathbf{x}, \mathbf{v}_{max} \rangle = \frac{1}{\sqrt{2}}(x_1 + x_2)$$

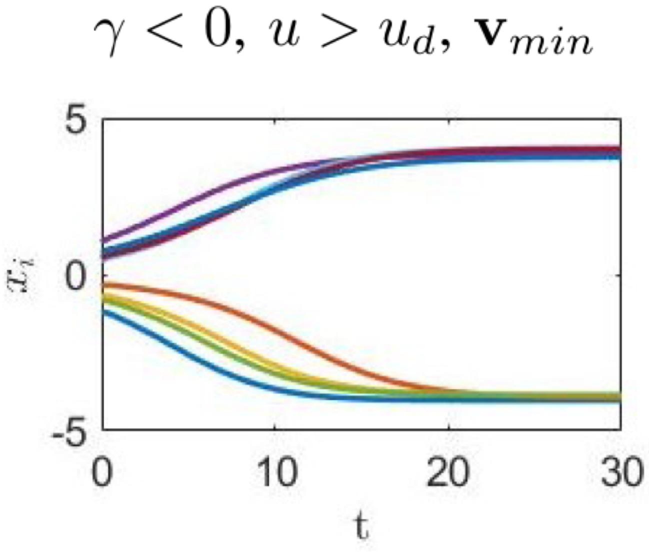
$$\langle \mathbf{b}, \mathbf{v}_{max} \rangle = \frac{1}{\sqrt{2}}(b_1 + b_2)$$

Bizyaeva, Sorochkin, Franci, Leonard, “Control of agreement and disagreement cascades with distributed inputs,” arXiv:2103.14764, March 2021

Theorem 2
Opinion Formation as Pitchfork Bifurcation

The agreement and disagreement bifurcations in Theorem 1 are supercritical pitchfork bifurcations.





Example:

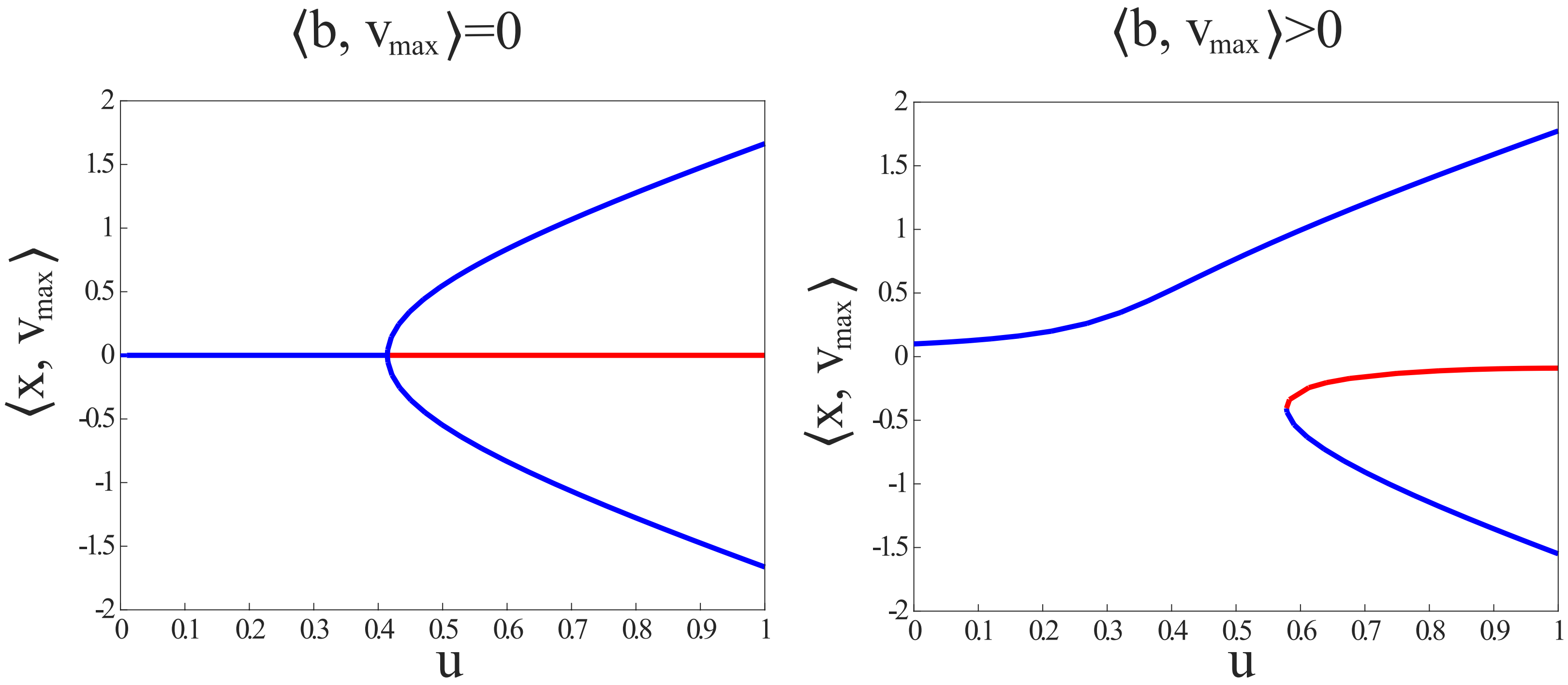
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \lambda_{min} = -1 \quad \mathbf{v}_{min} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\langle \mathbf{x}, \mathbf{v}_{min} \rangle = \frac{1}{\sqrt{2}}(x_1 - x_2)$$

$$\langle \mathbf{b}, \mathbf{v}_{min} \rangle = \frac{1}{\sqrt{2}}(b_1 - b_2)$$

Bizyaeva, Sorochkin, Franci, Leonard, “Control of agreement and disagreement cascades with distributed inputs,” arXiv:2103.14764, March 2021

Theorem 3
Eigenvector centrality determines influence of input on solution



Symmetric pitchfork bifurcation and its unfolding in agreement regime with 3 agents communicating over undirected line graph. Blue (red) curves are stable (unstable) equilibria. Vertical axis is projection of equilibria onto $W(\lambda_{max})$.

Parameters: $d = \alpha = \gamma = 1$. Left: $\mathbf{b} = (0.05, 0, -0.05)$; right: $\mathbf{b} = 0.1\mathbf{v}_{max} + (0.05, 0, -0.05)$.

Bizyaeva, Sorochkin, Franci, Leonard, “Control of agreement and disagreement cascades with distributed inputs,” arXiv:2103.14764, March 2021

Feedback dynamics for attention parameters

State feedback dynamics for u_i for each agent i to track saturated norm of its observation of opinion of system:

$$\tau_u \frac{du_i}{dt} = -u_i + S_u \left(x_i^2 + \sum_{k=1}^{N_a} (a_{ik} x_k)^2 \right).$$

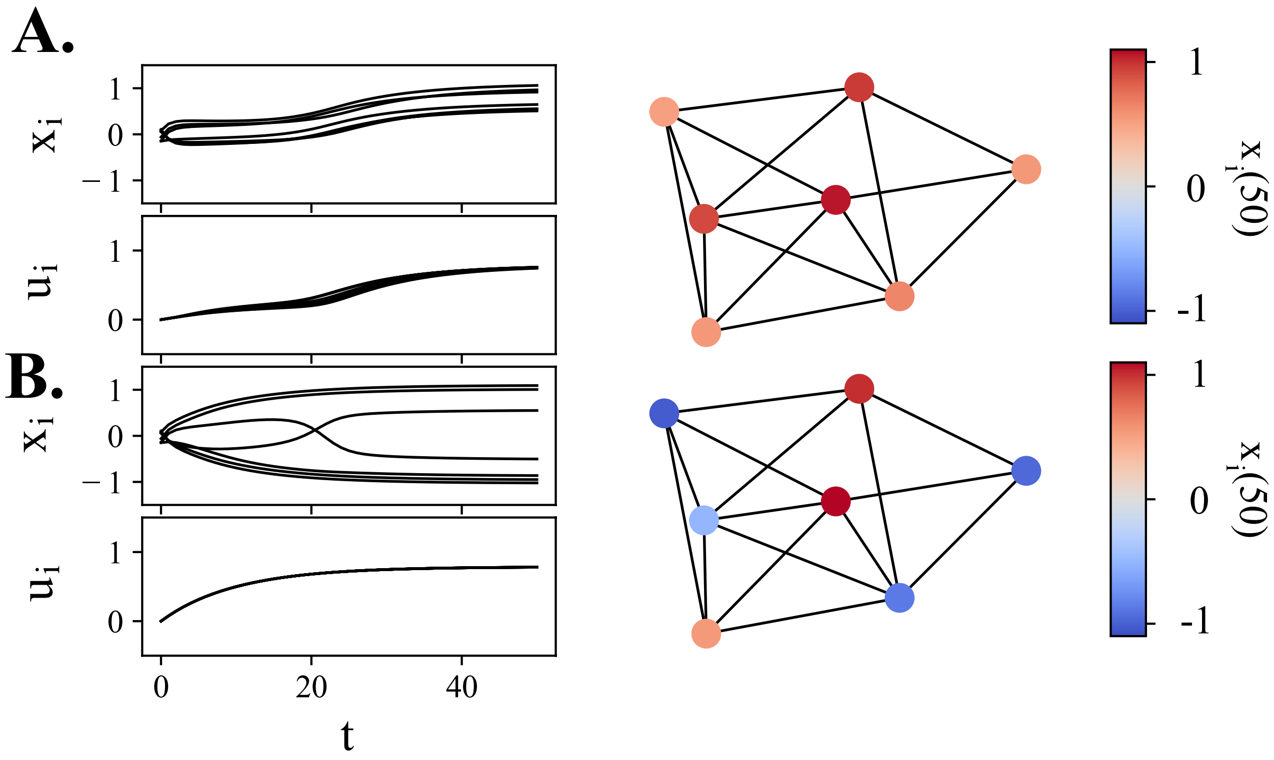
S_u takes the form of the Hill activation function:

$$S_u(y) = u + (\bar{u} - u) \frac{y^n}{(y_{th})^n + y^n}, \quad y_{th} > 0$$

Control of cascade

Input magnitude $\|\mathbf{b}\|$ and relative orientation $\mathbf{b}\angle\mathbf{v}_c := \langle \mathbf{v}_c, \mathbf{b}\rangle/\|\mathbf{b}\|$ control trigger of network opinion cascade.

Threshold for cascade depends on y_{th}. $\mathbf{v}_c$ is $\mathbf{v}_{max}$ or $\mathbf{v}_{max}$ depending on $\lambda > 0$ or $\lambda < 0$.



A. Agreement cascade, $\gamma = 1$, $\underline{u} = u_a - 0.01$, $\bar{u} = u_a + 0.6$; B. Disagreement cascade, $\gamma = -1$, $\underline{u} = u_d - 0.01$, $\bar{u} = u_d + 0.6$.

Parameters: $d = 1$, $n = 3$, $u_{th} = 0.4$, $\tau_u = 10$, $\alpha = 1$, $d = 1$. Each $x_i(0) \in \mathcal{N}(0, 0.1)$; $u_i(0) = 0$; $b_i \in \mathcal{N}(0, 0.2)$.

Bizyaeva, Sorochkin, Franci, Leonard, “Control of agreement and disagreement cascades with distributed inputs,” arXiv:2103.14764, March 2021

Application: Learning and cooperation in multi-agent finite games

Set of N_a agents (players): $\{1, \ldots, N_a\}$

Each agent has finite set of actions (pure strategies): $\{1, \ldots, N_o\}$

Mixed strategy of agent i is probability distribution over actions: $X_i = (x_{i1}, \ldots, x_{iN_o}) \in \Delta$

Mixed strategy profile of set of agents: $X = (X_1, \ldots, X_{N_a}) \in \Delta^{N_a}$

Payoff to agent i for selection of strategy j: $U_{ij}(X)$

Expected payoff to agent i for mixed strategy X_i: $U_i(X) = \sum_{j=1}^{N_o} x_{ij} U_{ij}(X)$

$$\Delta = \{y \in \mathbb{R}^{N_o}_{\geq 0} \mid \|y\|_1 = 1\}$$

Exponentially Discounted Reinforcement Learning (EXP-D-RL)

Score-based RL scheme modeled in continuous time (repeat indefinitely with infinitesimal time step):

1) **Assessment Stage**: Each agent i keeps score $\hat{Z}_i = (\hat{z}_{i1}, \ldots, \hat{z}_{iN_o}) \in \mathbb{R}^{N_o}$ based on received payoff

$$\frac{d\hat{z}_{ij}}{dt} = d\left(r_{ij} - \hat{z}_{ij}\right), \quad \hat{z}_{ij}(0) \in \mathbb{R}$$

for all $j = 1, \ldots, N_o$, where $r_{ij}(t) = U_{ij}(X(t))$, $d > 0$ is learning rate, and $\hat{z}_{ij}(0)$ is initial bias toward strategy i

2) **Choice Stage**: Each agent i maps its score $\hat{Z}_i$ into a mixed strategy $X_i \in \Delta$

$$x_{ij} = \frac{\exp\left(\eta^{-1}\hat{z}_{ij}\right)}{\sum_{l=1}^{N_o} \exp\left(\eta^{-1}\hat{z}_{il}\right)}, \quad \eta > 0$$

3) **Game Stage**: Each agent i plays game according to X_i

Gao & Pavel, IEEE TAC, 2021; Coucheney, Gaujal, & Mertikopoulos, 2015; Laraki and Mertikopoulos, 2013; Mertikopoulous & Sandholm, 2016

Example: 2-agent Prisoner's Dilemma

Strategy 1 is to *cooperate*

Strategy 2 is to *defect*

Probability that Agent i selects Strategy j:

$$x_{ij} = \frac{\exp(\hat{z}_{ij})}{\exp(\hat{z}_{i1}) + \exp(\hat{z}_{i2})} = \begin{cases} \frac{1}{1+\exp(\hat{z}_{i2}-\hat{z}_{i1})} & \text{if } j = 1 \\ \frac{1}{1+\exp(\hat{z}_{i1}-\hat{z}_{i2})} & \text{if } j = 2 \end{cases}$$

Reward assigned to Agent i for selecting Strategy j:

$$r_{ij} = U_{ij}(x_{11}, x_{12}, x_{21}, x_{22})$$

Example: 2-agent Prisoner's Dilemma

- Agent 1:

$$\begin{aligned}\frac{d\hat{z}_{11}}{dt} &= -d\hat{z}_{11} + dr_{11}\\ \frac{d\hat{z}_{12}}{dt} &= -d\hat{z}_{12} + dr_{12}\end{aligned}$$

  - Reward $r_1 = (r_{11}, r_{12})$:

$$\begin{pmatrix} r_{11} \\ r_{12} \end{pmatrix} = \begin{pmatrix} r_a & r_b \\ r_c & r_d \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}$$

- Agent 2:

$$\begin{aligned}\frac{d\hat{z}_{21}}{dt} &= -d\hat{z}_{21} + dr_{21}\\ \frac{d\hat{z}_{22}}{dt} &= -d\hat{z}_{22} + dr_{22}\end{aligned}$$

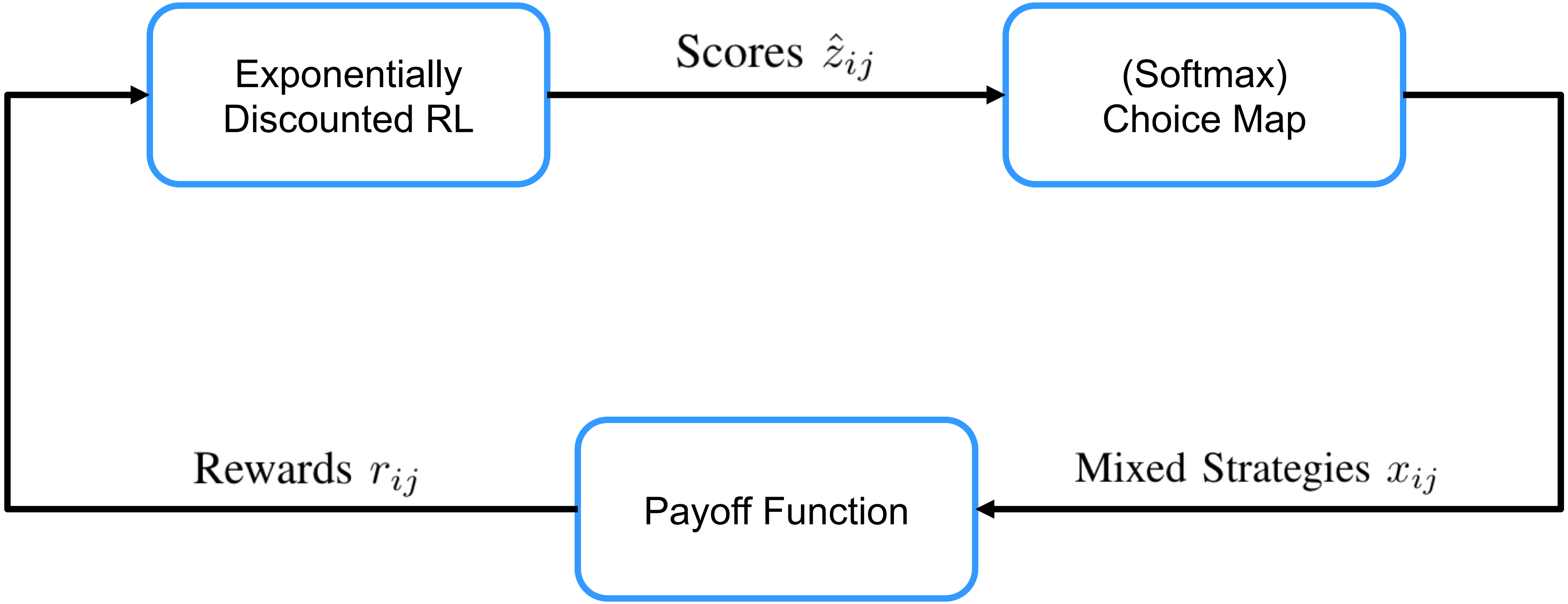
  - Reward $r_2 = (r_{21}, r_{22})$:

$$\begin{pmatrix} r_{21} \\ r_{22} \end{pmatrix} = \begin{pmatrix} r_a & r_b \\ r_c & r_d \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$$

$$r_c > r_a > r_d > r_b \text{ and } r_a + r_d = r_b + r_c$$

Nash equilibrium in which both agents defect is only stable solution

Exponentially Discounted Reinforcement Learning (EXP-D-RL)



Exponentially Discounted RL

Scores $\hat{z}_{ij}$

(Softmax) Choice Map

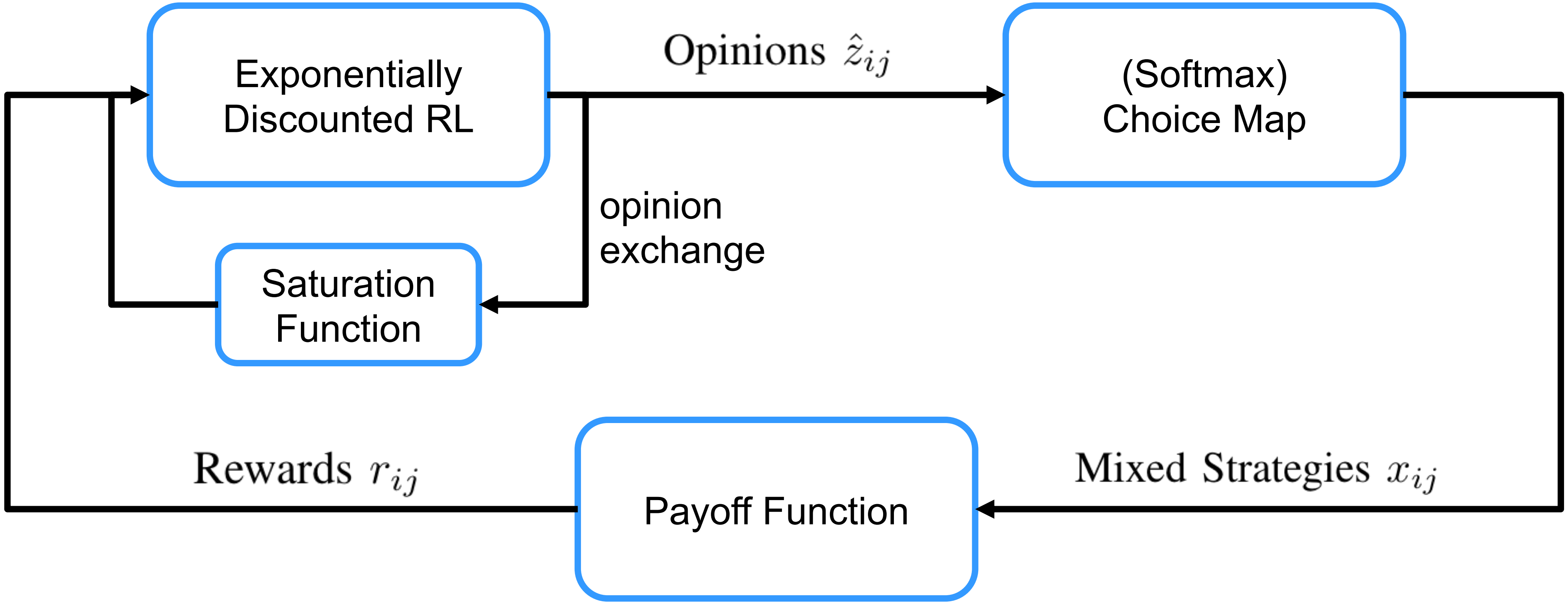
Mixed Strategies x_{ij}

Payoff Function

Rewards r_{ij}

Gao & Pavel, IEEE TAC, 2021

Feedback Saturation Function of Opinions (Scores) Communicated over Network



Exponentially Discounted RL

Opinions $\hat{z}_{ij}$

(Softmax) Choice Map

opinion exchange

Saturation Function

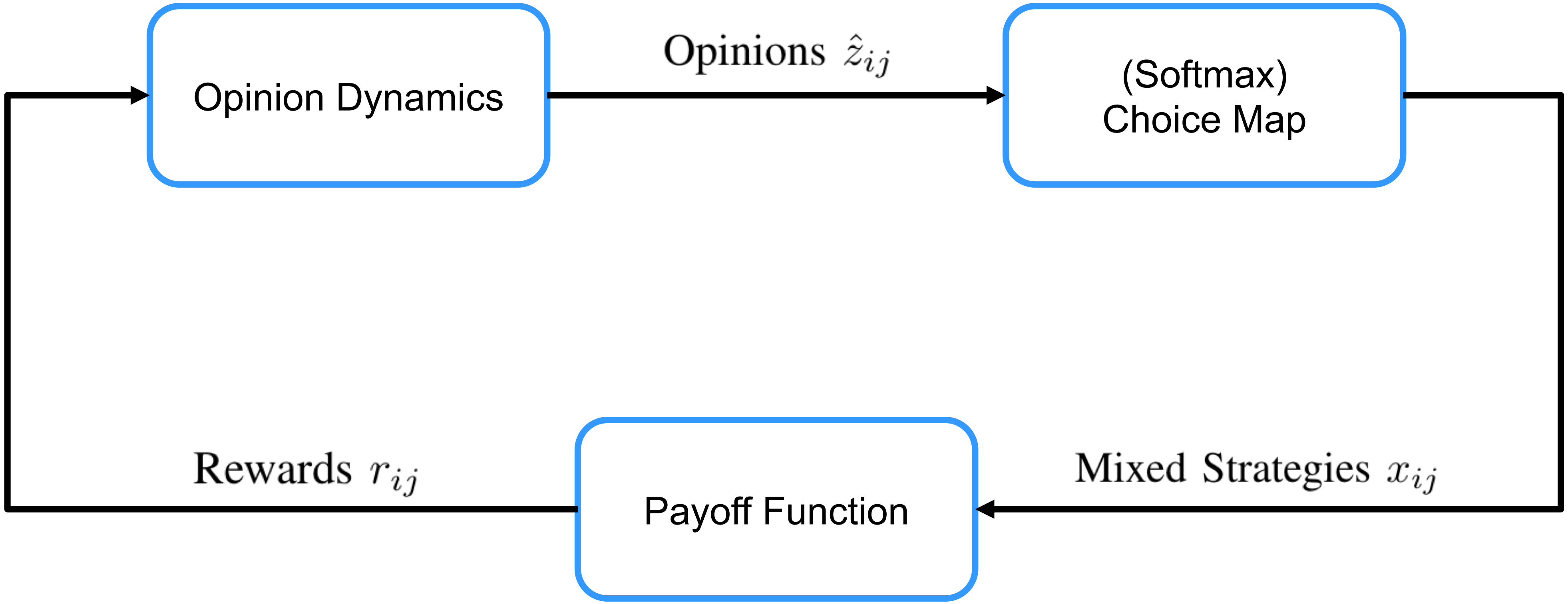
Rewards r_{ij}

Payoff Function

Mixed Strategies x_{ij}

Bizyaeva, Franci, Leonard, “A general model of opinion dynamics”, arXiv:2009.04332, Oct 2020

Opinion Dynamics Model


Opinion Dynamics

Opinions $\hat{z}_{ij}$

(Softmax)
Choice Map

Mixed Strategies x_{ij}

Payoff Function

Rewards r_{ij}

Bizyaeva, Franci, Leonard, “A general model of opinion dynamics”, arXiv:2009.04332, Oct 2020

2-agent Prisoner's Dilemma

$$z_{ij} = \hat{z}_{ij} - \sum_{l \neq j}^{N_o} \hat{z}_{il}$$

- Agent 1:

$$\frac{d\hat{z}_{11}}{dt} = -d\hat{z}_{11} + u\left(\tanh(\gamma z_{21})\right) + r_{11}$$
$$\frac{d\hat{z}_{12}}{dt} = -d\hat{z}_{12} + u\left(\tanh(\gamma z_{22})\right) + r_{12}$$

- Agent 2:

$$\frac{d\hat{z}_{21}}{dt} = -d\hat{z}_{21} + u\left(\tanh(\gamma z_{11})\right) + r_{21}$$
$$\frac{d\hat{z}_{22}}{dt} = -d\hat{z}_{22} + u\left(\tanh(\gamma z_{12})\right) + r_{22}$$

– Reward $r_1 = (r_{11}, r_{12})$:

$$\begin{pmatrix} r_{11} \\ r_{12} \end{pmatrix} = \begin{pmatrix} r_a & r_b \\ r_c & r_d \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}$$

– Reward $r_2 = (r_{21}, r_{22})$:

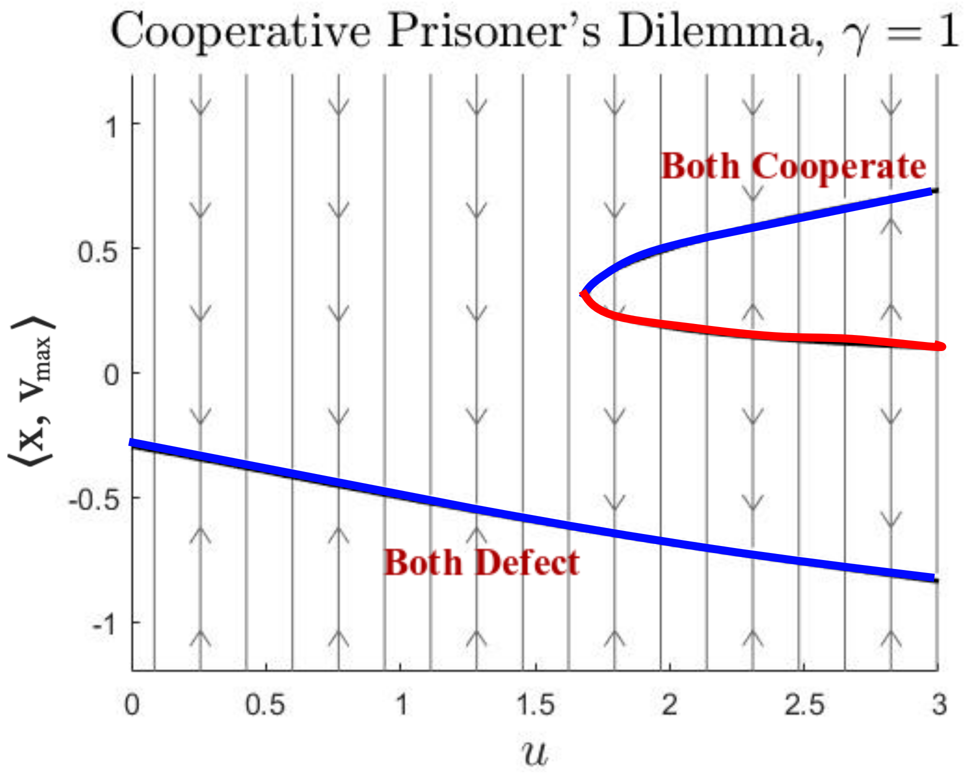
$$\begin{pmatrix} r_{21} \\ r_{22} \end{pmatrix} = \begin{pmatrix} r_a & r_b \\ r_c & r_d \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$$

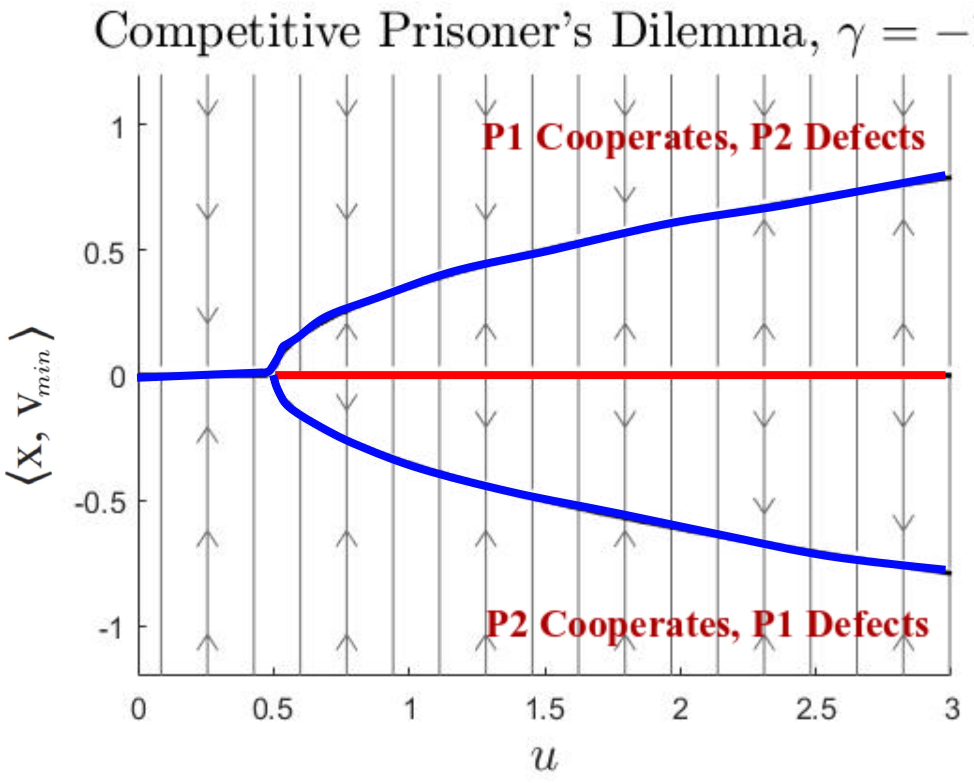
$$r_c > r_a > r_d > r_b \text{ and } r_a + r_d = r_b + r_c$$

Nash equilibrium in which both agents defect is **NOT** only stable solution

Bizyaeva, Park, Franci, Leonard, "Opinion dynamics and learning to cooperate in multi-agent games," in preparation

2-agent Prisoner's Dilemma





Both defect (Nash) and **Both cooperate**

are bistable equilibria for high enough u

p controls bifurcation point where $p = r_a - r_b + r_d - r_c$

2-agent Prisoner's Dilemma

$$z_{ij} = \hat{z}_{ij} - \sum_{l \neq j}^{N_o} \hat{z}_{il}$$

- Agent 1:

$$\frac{d\hat{z}_{11}}{dt} = -d\hat{z}_{11} + u\left(\tanh(\gamma z_{21})\right) + r_{11}$$
$$\frac{d\hat{z}_{12}}{dt} = -d\hat{z}_{12} + u\left(\tanh(\gamma z_{22})\right) + r_{12}$$

- Agent 2:

$$\frac{d\hat{z}_{21}}{dt} = -d\hat{z}_{21} + u\left(\tanh(\gamma z_{11})\right) + r_{21}$$
$$\frac{d\hat{z}_{22}}{dt} = -d\hat{z}_{22} + u\left(\tanh(\gamma z_{12})\right) + r_{22}$$

– Reward $r_1 = (r_{11}, r_{12})$:

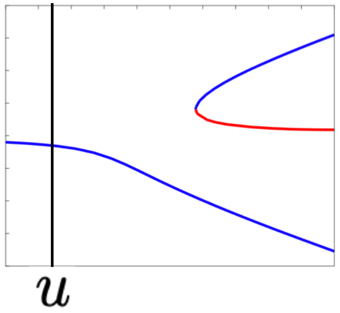
$$r_1 = \begin{pmatrix} r_{11} \\ r_{12} \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}$$

– Reward $r_2 = (r_{21}, r_{22})$:

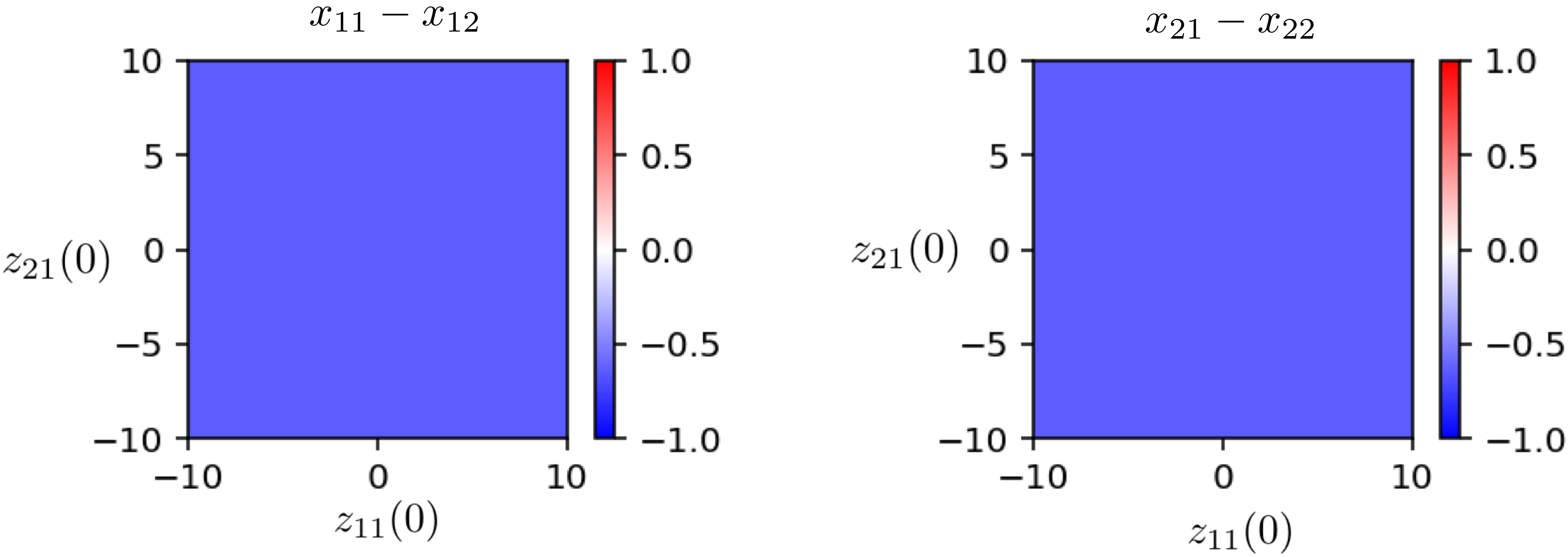
$$r_2 = \begin{pmatrix} r_{21} \\ r_{22} \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$$

Bizyaeva, Park, Franci, Leonard, "Opinion dynamics and learning to cooperate in multi-agent games," in preparation

2-agent Prisoner's Dilemma

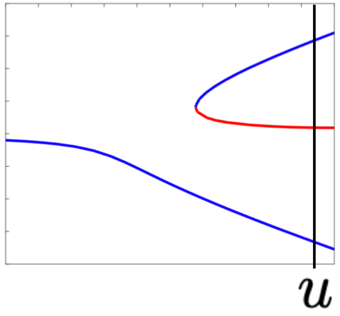


- parameter selection: $d = 1$, $u = 0.2$, $\gamma = 1$

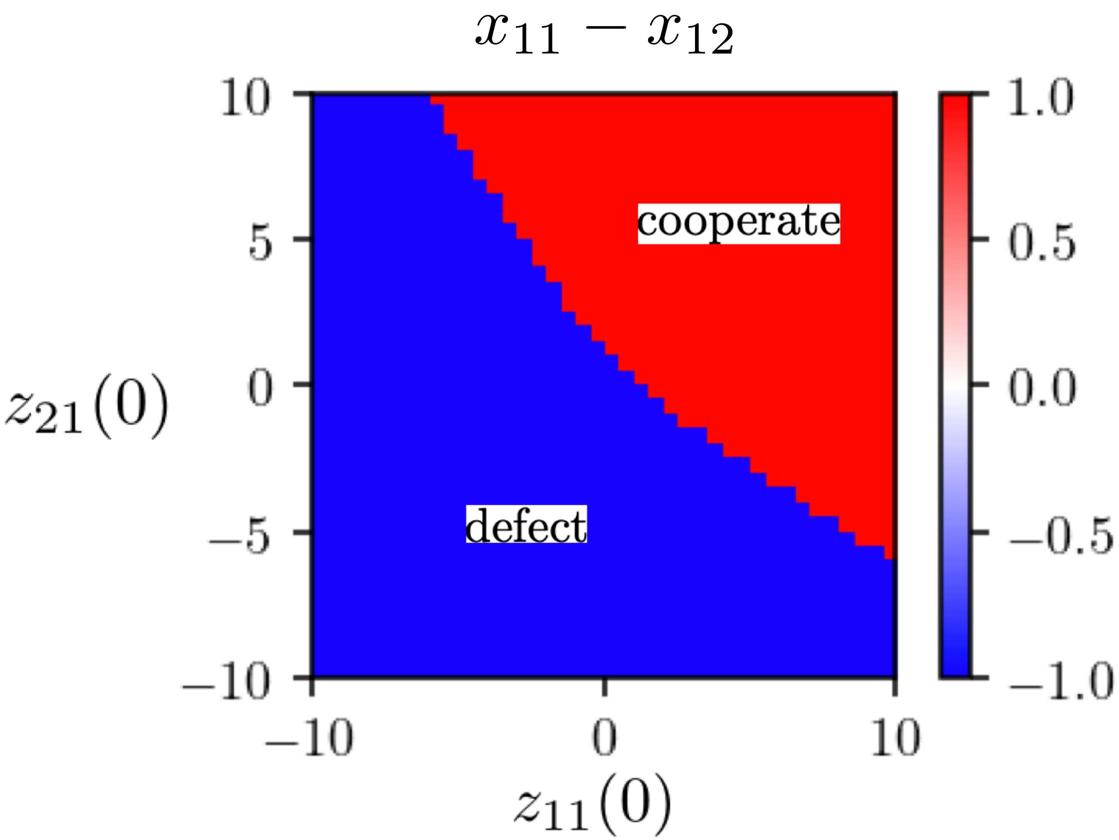


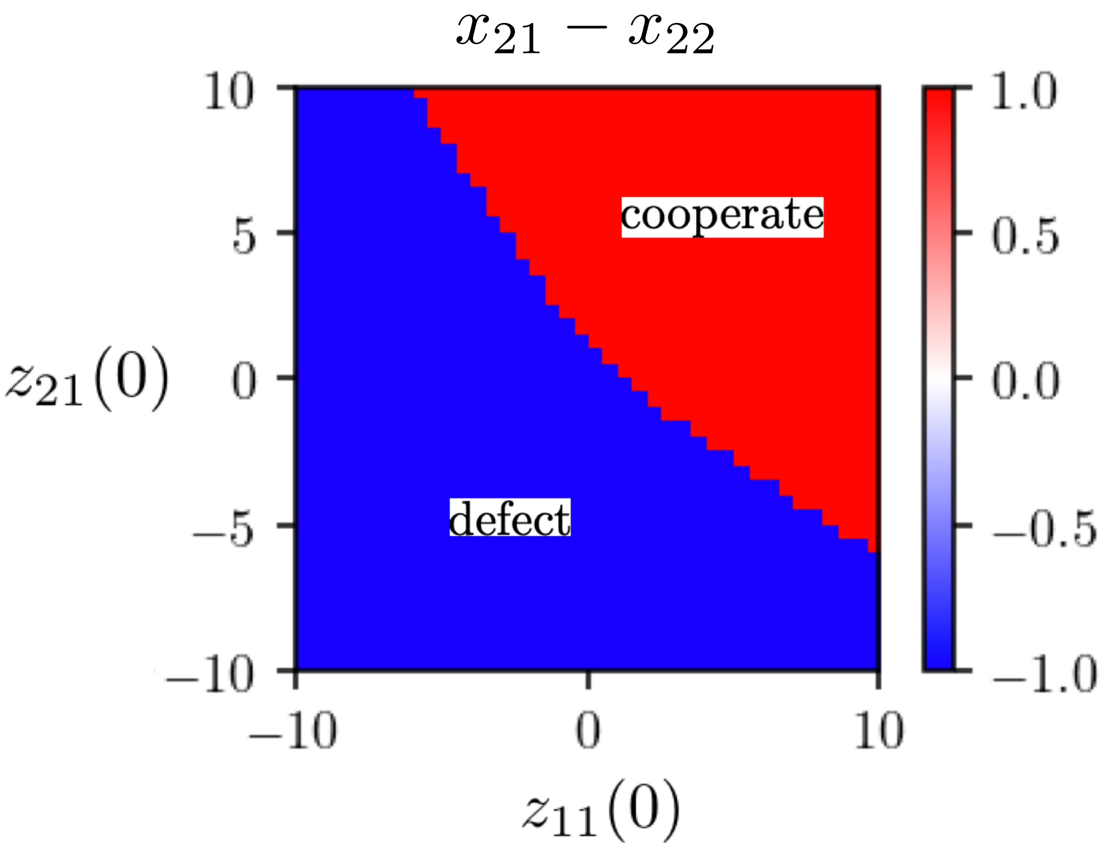
Bizyaeva, Park, Franci, Leonard, "Opinion dynamics and learning to cooperate in multi-agent games," in preparation

2-agent Prisoner's Dilemma



- parameter selection: $d = 1$, $u = 3$, $\gamma = 1$





Bizyaeva, Park, Franci, Leonard, "Opinion dynamics and learning to cooperate in multi-agent games," in preparation

2-agent Prisoner's Dilemma with attention dynamics

- Agent 1:

$$
\begin{aligned}
\frac{d\hat{z}_{11}}{dt} &= -d\hat{z}_{11} + u_1\left(\tanh(\gamma z_{21})\right) + r_{11} \\
\frac{d\hat{z}_{12}}{dt} &= -d\hat{z}_{12} + u_1\left(\tanh(\gamma z_{22})\right) + r_{12} \\
\frac{du_1}{dt} &= -u_1 + \frac{5}{2}\left(\tanh(x_{21}r_{11} + x_{22}r_{12}) + 1\right)
\end{aligned}
$$

  - Reward $r_1 = (r_{11}, r_{12})$:

$$
r_1 = \begin{pmatrix} r_{11} \\ r_{12} \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}
$$

- Agent 2:

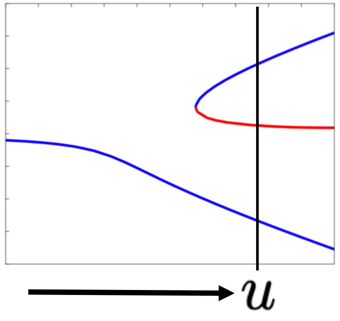
$$
\begin{aligned}
\frac{d\hat{z}_{21}}{dt} &= -d\hat{z}_{21} + u_2\left(\tanh(\gamma z_{11})\right) + r_{21} \\
\frac{d\hat{z}_{22}}{dt} &= -d\hat{z}_{22} + u_2\left(\tanh(\gamma z_{12})\right) + r_{22} \\
\frac{du_2}{dt} &= -u_2 + \frac{5}{2}\left(\tanh(x_{11}r_{21} + x_{12}r_{22}) + 1\right)
\end{aligned}
$$

  - Reward $r_2 = (r_{21}, r_{22})$:

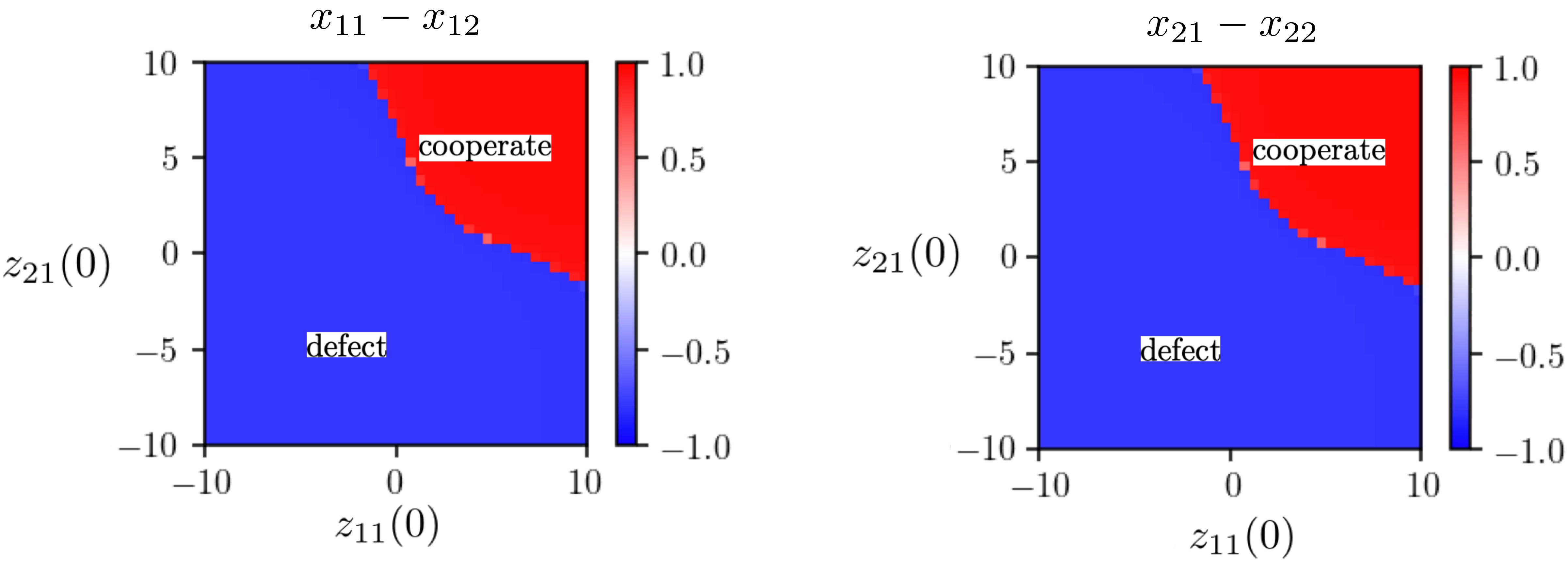
$$
r_2 = \begin{pmatrix} r_{21} \\ r_{22} \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}
$$

Bizyaeva, Park, Franci, Leonard, "Opinion dynamics and learning to cooperate in multi-agent games," in preparation

2-agent Prisoner’s Dilemma



- parameter selection: $d = 1,\ \gamma = 1$



Bizyaeva, Park, Franci, Leonard, “Opinion dynamics and learning to cooperate in multi-agent games,” in preparation

Application: Dynamic task allocation of a robot swarm

N_a is number of robots, N_o is number of tasks

$\mathcal{N}_i$ is set of robot i's neighbors

$\mu_j \in [0, 1]$, where $\sum_{j=1}^{N_o} \mu_j = 1$, is the priority of task j

ν_{ij} is the intrinsic zealousness of robot i to perform task j

Let $b_{ij} = u\mu_j(|\mathcal{N}_i| + \nu_{ij})$, $\alpha = 0$, $\tilde{\gamma} = -\gamma > 0$, then

$$\dot{z}_{ij} = F_{ij}(\mathbf{Z}) - \frac{1}{N_o} \sum_{l=1}^{N_o} F_{il}(\mathbf{Z})$$

$$F_{ij}(\mathbf{Z}) = -z_{ij} + u \left(\mu_j \left(|\mathcal{N}_i| + \nu_{ij} \right) - \frac{1}{2} \sum_{k \in \mathcal{N}_i} S\left(2\tilde{\gamma} z_{kj} \right) \right)$$

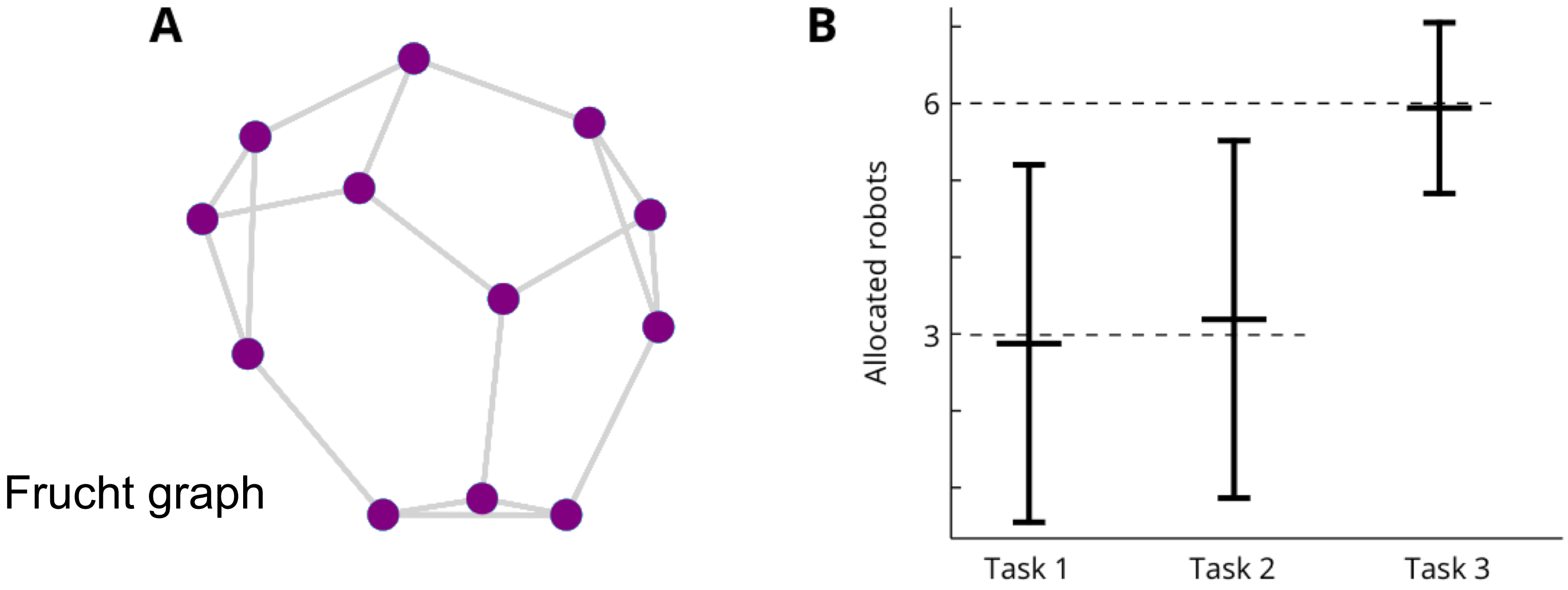
$\nu_{ij} = 0$: robot i updates preference for task j by $\mu_j \, |\mathcal{N}_i| - N_i^j$

$\nu_{ij} > 0$: robot i has greater tendency i to choose task j

where N_i^j is # of robots in $\mathcal{N}_i$ such that $z_{kj} > 0$

Franzi, Bizyaeva, Park, Leonard, "Analysis and control of agreement and disagreement cascades," *Swarm Intelligence,* 2021

12 Robots Self-allocating Across 3 Tasks



$$|\mathcal{N}_i| = 3,\ \mu_1 = \mu_2 = 0.3,\ \mu_3 = 0.4,\ \tilde{\gamma} = 1.0,\ u = 2u_d$$

100 simulations with small random ν_{ij} and small random initial conditions

Franzi, Bizyaeva, Park, Leonard, "Analysis and control of agreement and disagreement cascades," *Swarm Intelligence,* 2021

Dynamics Task Allocation with Attention Dynamics

Zealous robot senses real-time changes in task urgency

Suppose one zealous robot i_z detects increase in urgency of Task 3 of magnitude ρ_3

Then $\nu_{i_z 3} = 3\rho_3$ and effective urgency of Task 3 perceived by robot i_z is $\mu_3 + \rho_3$

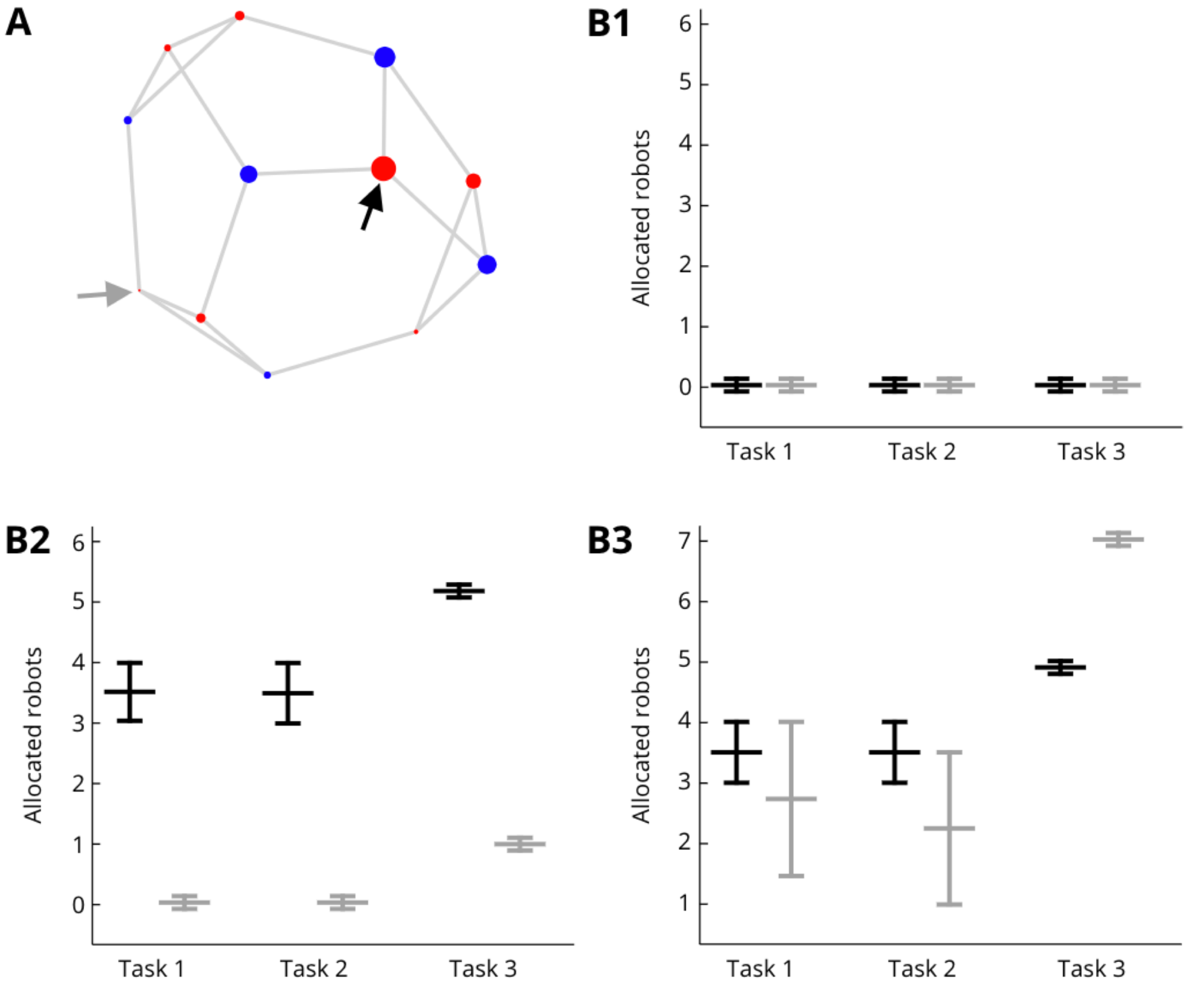
All agents have attention dynamics:

$$\tau_u \dot{u}_i = -u + u_{min} + (u_{max} - u_{min}) S_u(\|\boldsymbol{Z}_i\|), \quad S_u(y) = \frac{y^n}{u_{th}^n + y^n}$$

with $u_{min} = u_d/2$, $u_{max} = 2u_d$, $u_{th} = 0.1$, $n = 5$

Franzi, Bizyaeva, Park, Leonard, “Analysis and control of agreement and disagreement cascades,” *Swarm Intelligence,* 2021

Dynamics Task Allocation with Attention Dynamics: One Zealous Robot



Black: zealous robot is most (disagreement) central node

Gray: zealous robot is least (disagreement) central node

B1: $\rho_3 = 10^{-0.9}$, B2: $\rho_3 = 10^{-0.6}$, B3: $\rho_3 = 10^{-0.3}$

$|\mathcal{N}_i| = 3,\ \mu_1 = \mu_2 = 0.3,\ \mu_3 = 0.4,\ \tilde{\gamma} = 1.0,\ u = 2u_d$

Franzi, Bizyaeva, Park, Leonard, "Analysis and control of agreement and disagreement cascades," *Swarm Intelligence,* 2021

Final Remarks

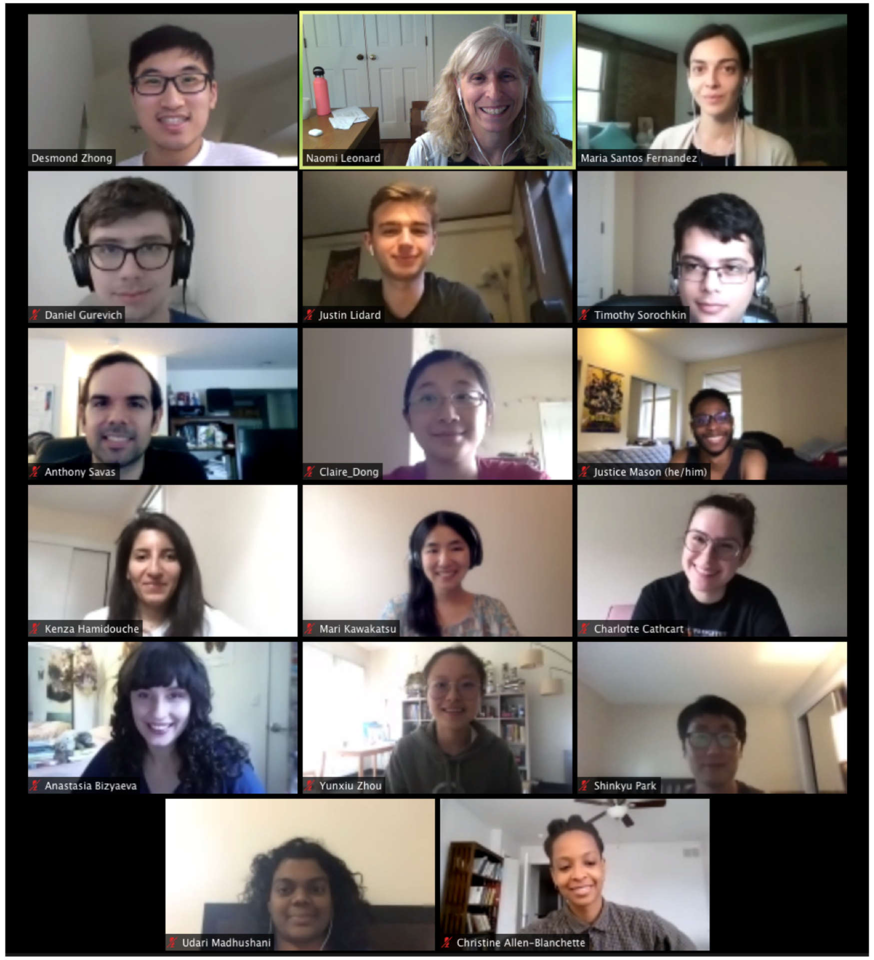
Nonlinear multi-agent, multi-option opinion dynamics for decision making and learning

- Naturally extends models that update based on weighted-average of neighbor opinions
- Quick, reliable, and informed decision-making in response to external cues:
    - opinions form through bifurcation
    - attention dynamics and controllable cascades
    - breaks deadlocks
    - tunable sensitivity
- Rich suite of behaviors: multi-stability of agreement and disagreement opinion configuration
- Analytical tractability: systematically leverage network structure in design

Applications illustrated:

- Reinforcement learning in multi-agent finite games
- Dynamic multi-robot task allocation

Thank you!



Research group

Tim, Amara, and Lily Leonard