Collective Reinforcement Learning Dynamics for Sustainability Economics

Workshop on Modern Applications of Control Theory and Reinforcement Learning | May 21, 2025

Wolfram Barfuss

Argelander Professor of Integrated Systems Modeling for Sustainability Transitions

zef
Zentrum für Entwicklungsforschung
Universität Bonn

TRA SUSTAINABLE FUTURES
UNIVERSITÄT BONN

Plot

CSS+MARL $\longrightarrow$ SusEcon

Are we smart enough for the good life?

The good life for all within a healthy planet

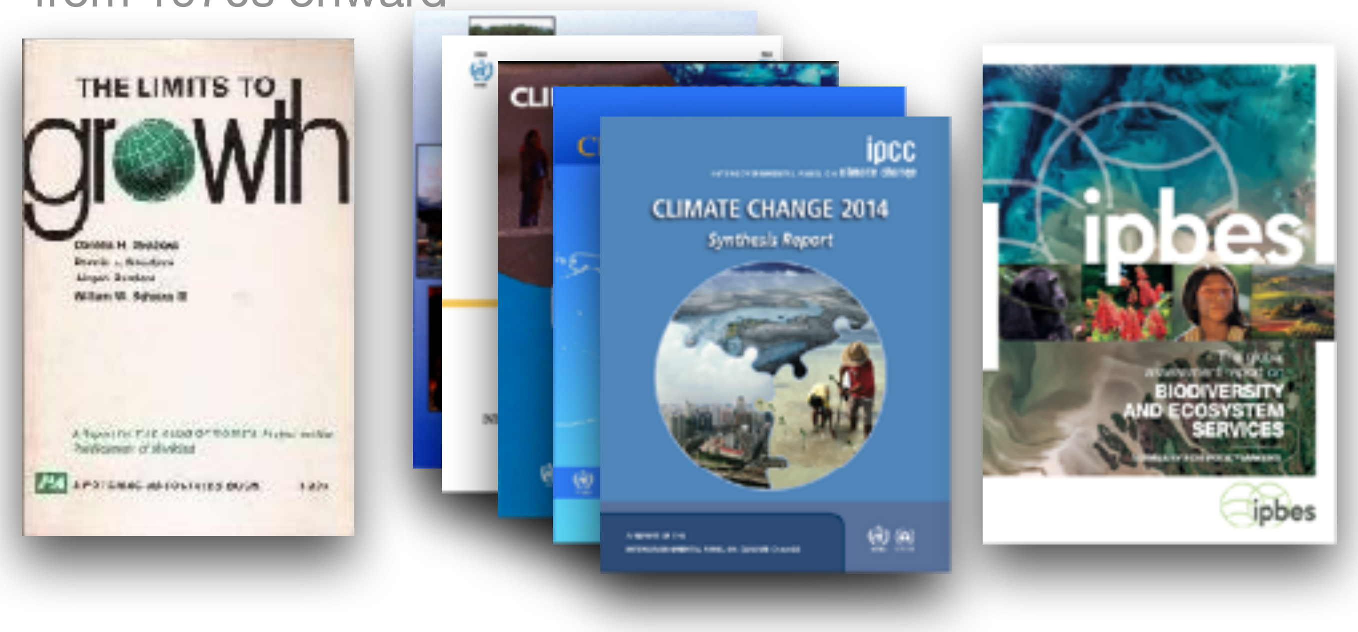
is the challenge of the 21st century

Azote for Stockholm Resilience Centre, based on analysis in Persson et al. (2022) and Richardson et al. (2023), Raworth (2012, 2017), Steffen et al. (2018)

Urgent & large-scale collective action

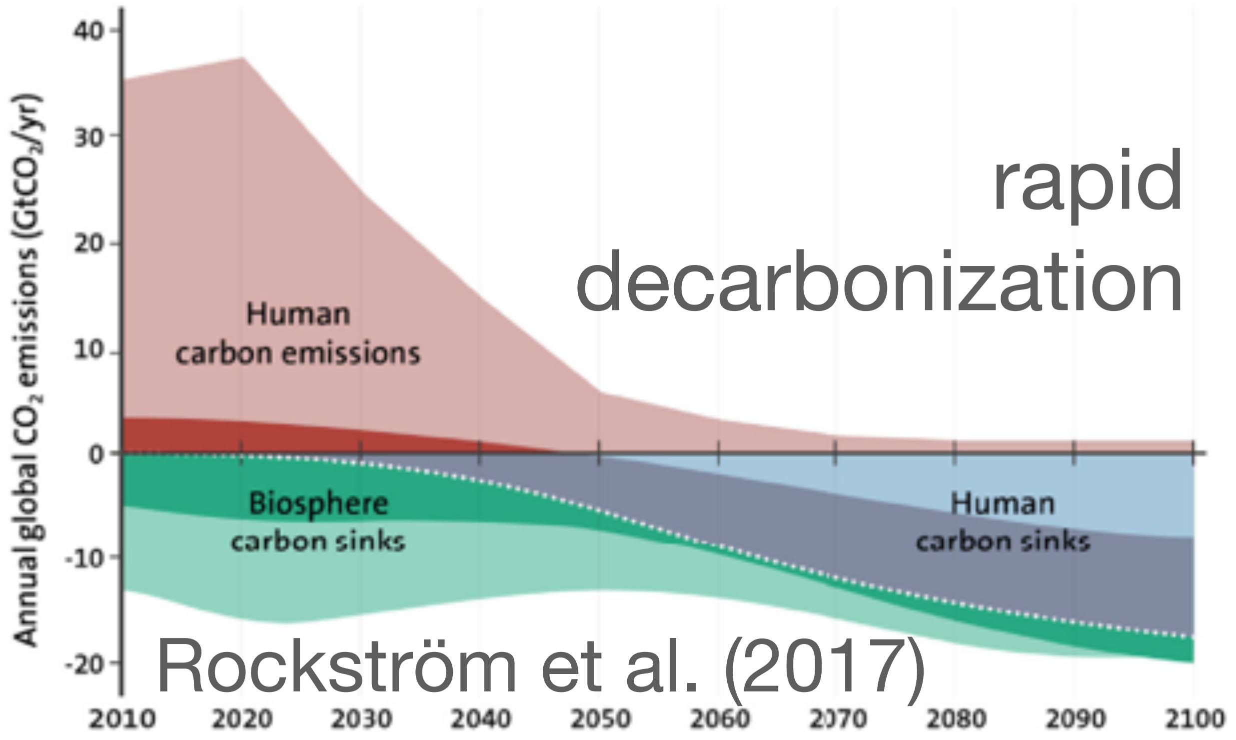
How to enter a safe and just space?

from 1970s onward



Report danger
(with more certainty)

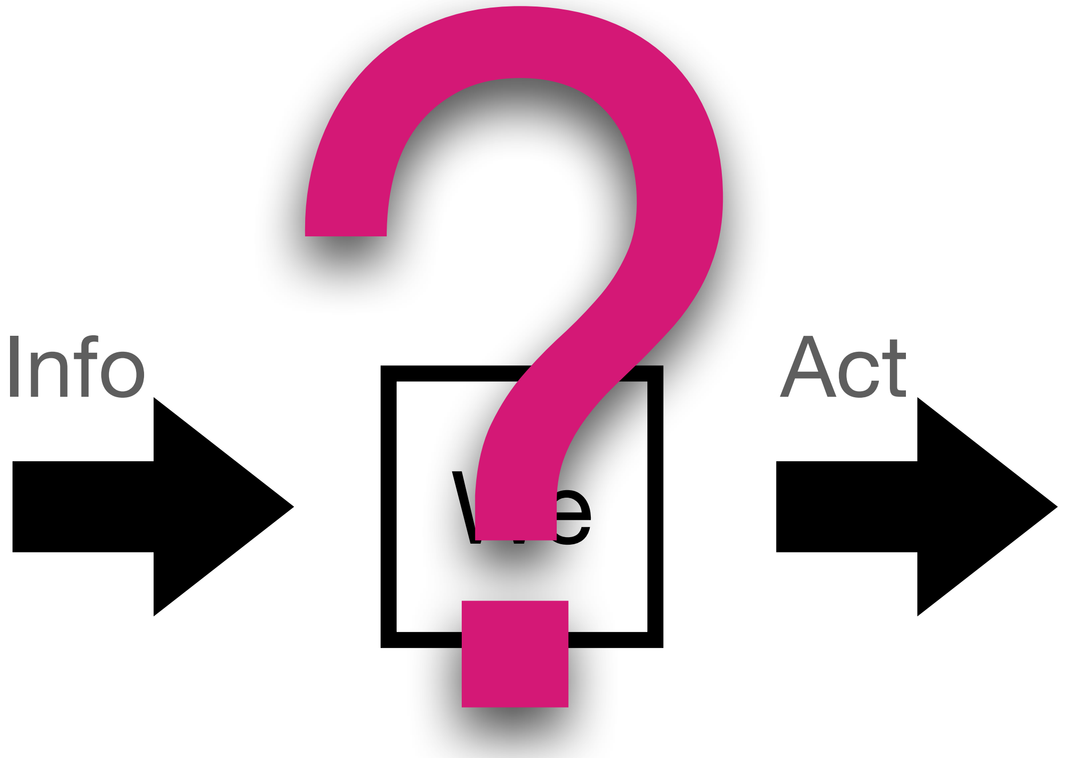
Technology
Social Media
Artificial Intelligence



Propose solutions

Technological
Institutional

We ***know***
there is a problem



We ***know***
what to do

Outline

Why | Collective action challenges in human-machine ecologies

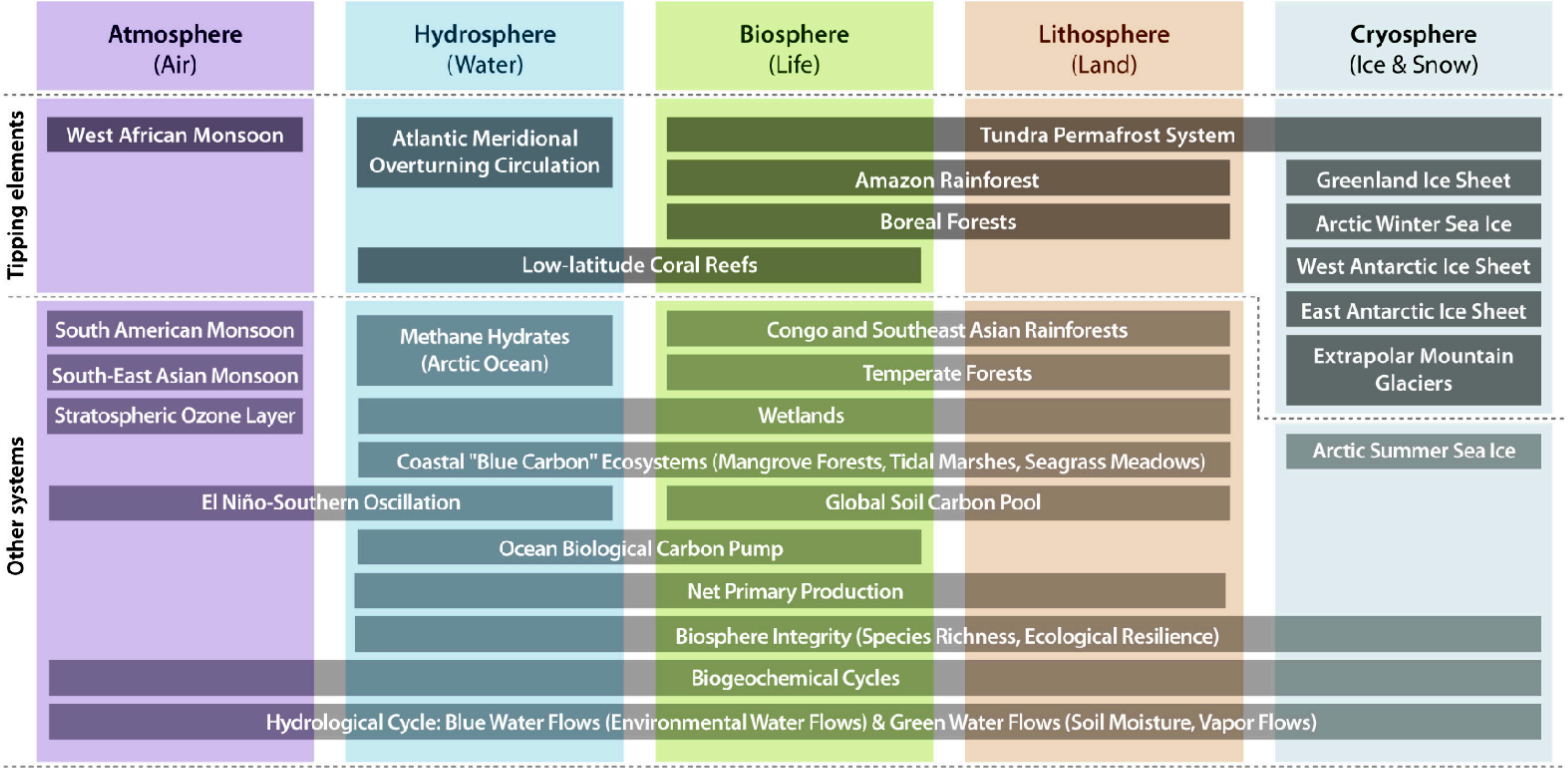
How | Building bridges

What | Emergent phenomena from cognition in contexts

WHY | Collective action challenges in human-machine ecologies

Planetary commons

Nature’s regulating and maintenance services requiring cooperation



Rockström et al. 2024 The planetary commons: A new paradigm for safeguarding Earth-regulating systems in the Anthropocene

The challenge of collective action | **Cooperation**

Individual
pays a cost *c*

→

Collective
All receive benefit *b*
with $b<c<Nb$

The tragedy of the commons

Hardin (1968)

A conflict of interest between individual and collective

Greed to exploit others and
fear of being exploited by others

Individual interest → defection

Everyone is better off cooperating

Collective interest → cooperation

	I cooperate	I defect
you cooperate	2b-c | 2b-c	-c | b
you defect	b | -c	0 | 0

Solution | Outside central authorities or privatization

Problem | Not possible for planetary commons

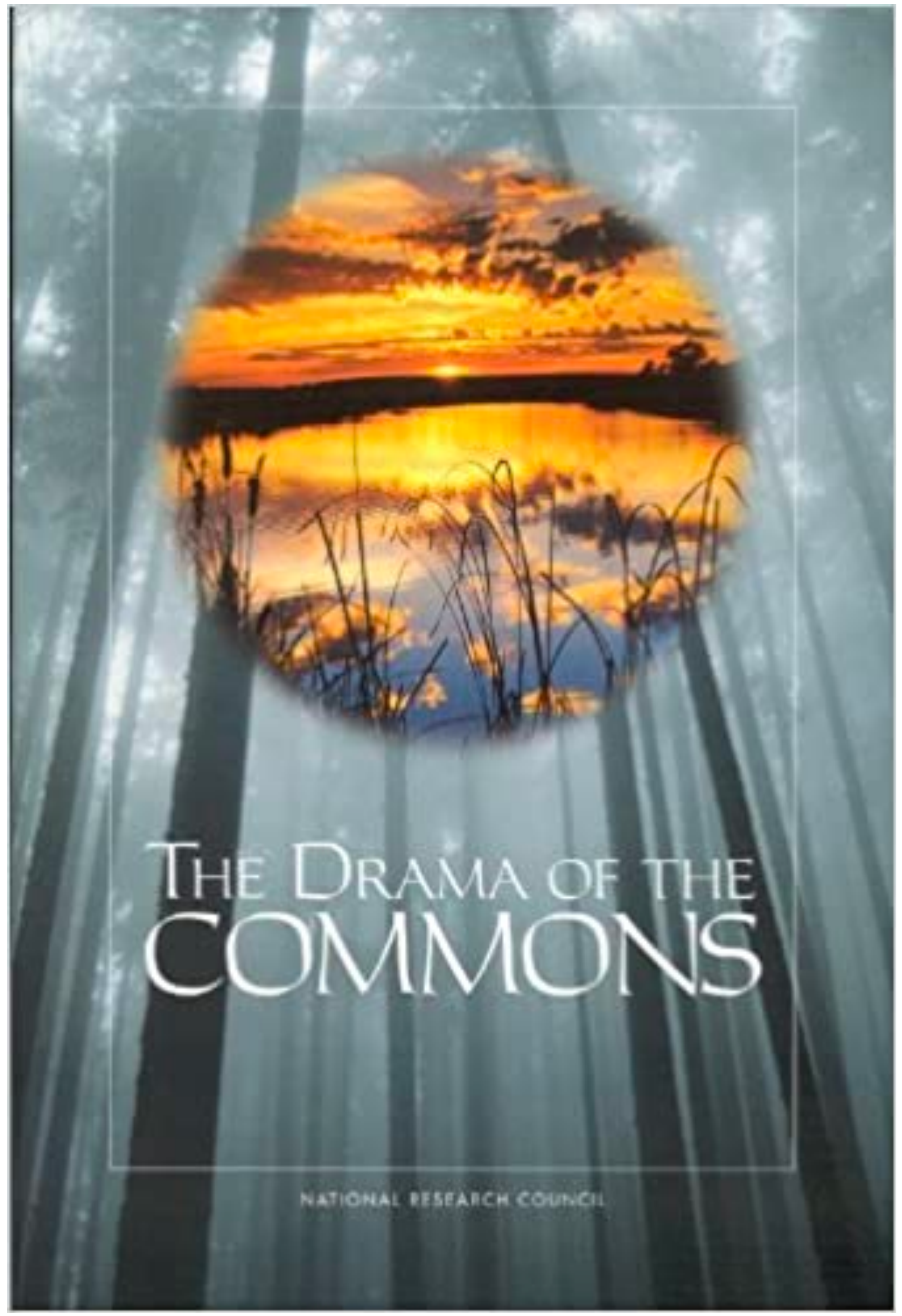
Governance of local commons

Tragedy is not inevitable - Things are **not as simple** as they seem in the prototypical model. **Human motivation** is complex […], and the **resource systems themselves have dynamics** that influence their response to human use.

Centralized authorities tend to overuse the commons as well.

Solution | Bottom-up approaches, decentralized governance, social reciprocity

Problem | Difficult to scale to the global level



Ostrom et al. (2002). The drama of the commons

Intelligent technology should also cooperate

Setting the agenda in research

Comment

A huddle at the 2017 United Nations Climate Change Conference, where attendees cooperated on mutually beneficial joint actions on climate.

Cooperative AI: machines must learn to find common ground

Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson & Thore Graepel

"To help humanity solve **fundamental problems of cooperation**, scientists need to reconceive **artificial intelligence as deeply social**"

Dafoe et al. 2021 Cooperative AI: machines must learn to find common ground

Cooperation challenges in human-machine ecologies

Why are we not cooperating more toward a sustainable future for all?

Collective Behavior 👥
Social context
Heterogeneity
Structure
Scale

Cognitive Agency 🧠
Perception
Valuation
Decision-making
Action

Good modeling practices 🗺️
Coherence | Transparency | Sparsity

Environmental context 🌳
Smooth changes
Abrupt changes
Uncertain changes
Delayed consequences

Dynamics 🚀
Nonlinearities & Feedbacks
Stability & Resilience
Transient evolution
Critical transitions

Personal synthesis based on Schill et al. (2019), Elsawah et al. (2020), Müller et al. (2020), Constantino et al. (2021), Levin & Xepapadeas (2021), Farahbakhsh et al. (2022), Giupponi et al. (2022)

HOW | Building bridges

How to model? | **Three types of models**

Dynamic-system models

Transparent
Dynamics
Collective behavior
? agency ?
Environmental contexts

Equilibrium-based models

Transparent
Static
Collective behavior
Hyper-rational agency
Environmental contexts

Model desiderata
Transparent
Dynamics

How to integrate?

Often opaque
Dynamic
Collective behavior
Arbitrary agency
Environmental contexts

Agent-based models

Building bridges between communities

Complex Systems Science

Insightful - produced rich understanding on how cooperation can emerge

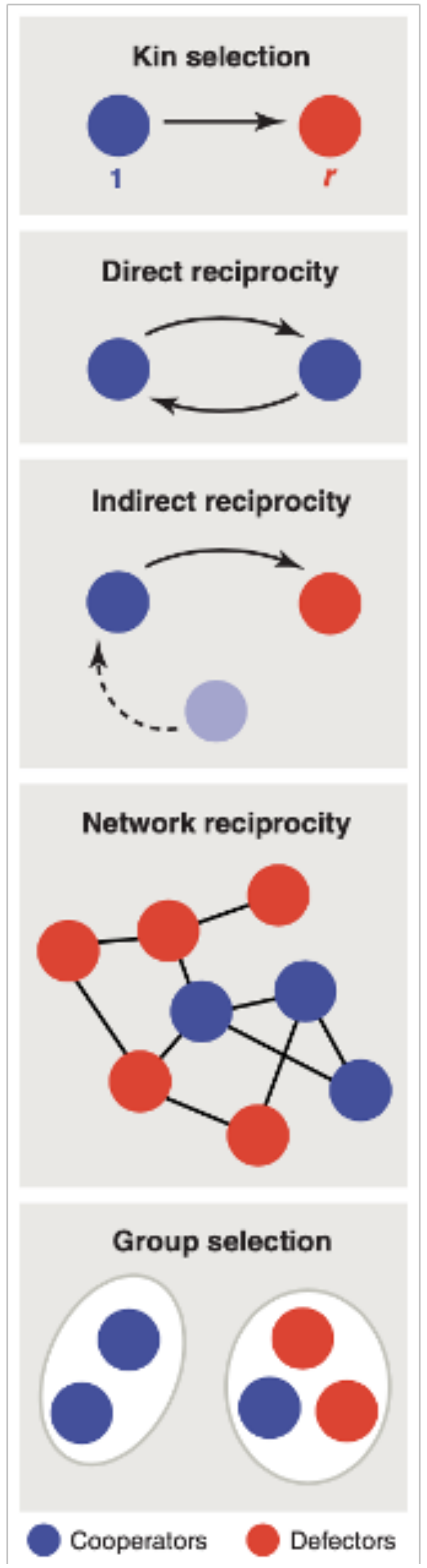
Simplistic - ignoring individual-level complexity and environmental context

Multiagent reinforcement learning

Rich - combining collectives of individually intelligent agents in changing environments

Obscure - highly stochastic, computationally expensive, and challenging to interpret

CSS Example | Five rules for the evolution of cooperation



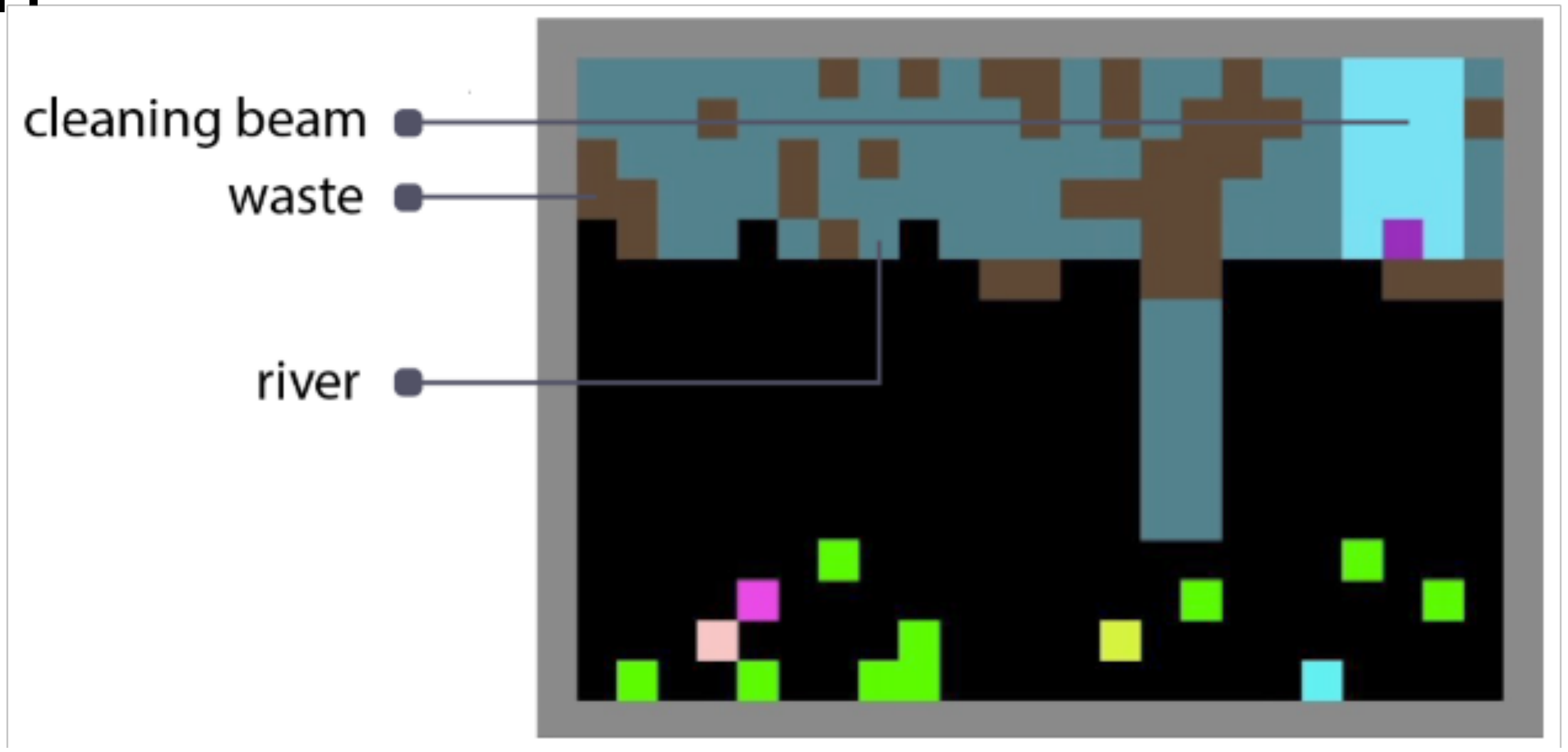
Cooperation as an emergent phenomenon, given …

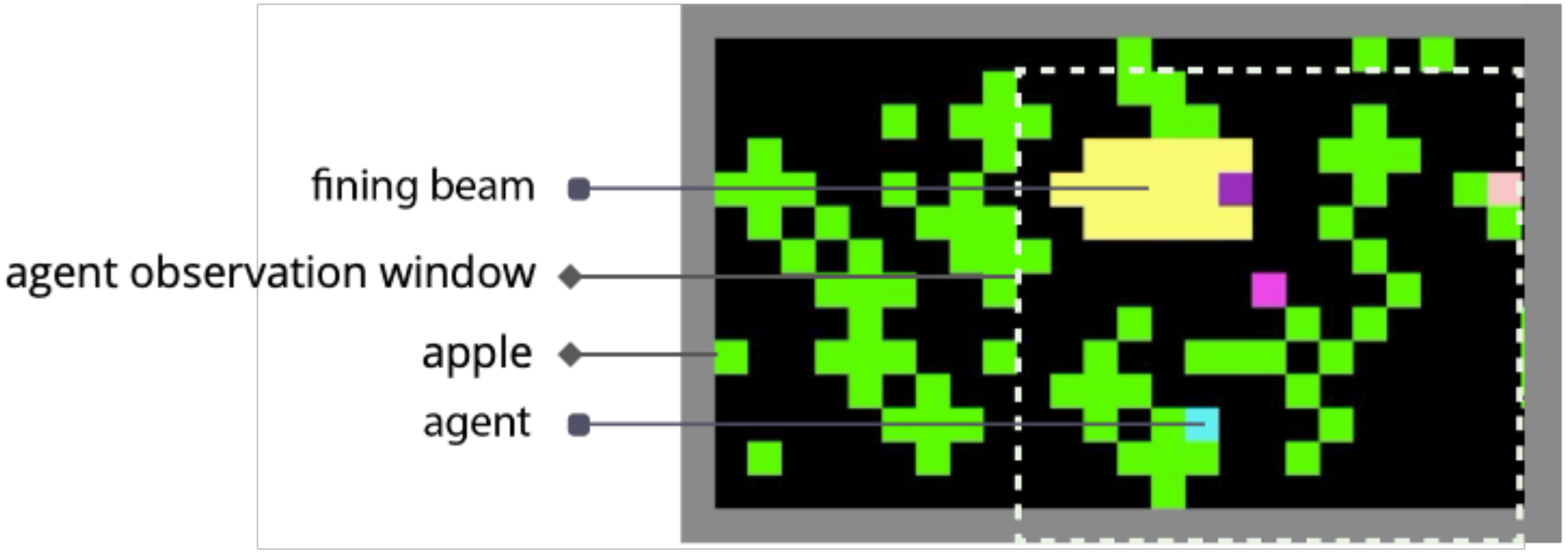
		Payoff matrix: C	Payoff matrix: D	Cooperation is… ESS	RD	AD	
Kin selection	C D	$(b-c)(1+r)$ $b-rc$	$br-c$ 0	$\frac{b}{c}>\frac{1}{r}$	$\frac{b}{c}>\frac{1}{r}$	$\frac{b}{c}>\frac{1}{r}$	*r*…genetic relatedness
Direct reciprocity	C D	$(b-c)/(1-w)$ b	$-c$ 0	$\frac{b}{c}>\frac{1}{w}$	$\frac{b}{c}>\frac{2-w}{w}$	$\frac{b}{c}>\frac{3-2w}{w}$	*w*…probability of next round
Indirect reciprocity	C D	$b-c$ $b(1-q)$	$-c(1-q)$ 0	$\frac{b}{c}>\frac{1}{q}$	$\frac{b}{c}>\frac{2-q}{q}$	$\frac{b}{c}>\frac{3-2q}{q}$	*q*…social acquaintanceship
Network reciprocity	C D	$b-c$ $b-H$	$H-c$ 0	$\frac{b}{c}>k$	$\frac{b}{c}>k$	$\frac{b}{c}>k$	*k*…number of neighbors
Group selection	C D	$(b-c)(m+n)$ bn	$(b-c)m-cn$ 0	$\frac{b}{c}>1+\frac{n}{m}$	$\frac{b}{c}>1+\frac{n}{m}$	$\frac{b}{c}>1+\frac{n}{m}$	*n*…group size *m*…number of groups

Nowak 2006 Five Rules for the Evolution of Cooperation

MARL Example | Cooperation in sequential social dilemmas

Cooperation not readily available as a single action
- must be learned as a sequence of action





Solution approaches

- Other-regarding preferences
- Other-influence
- Reputation and norms
- Contracts

Frontiers: LLM-based agents

Leibo et al. 2017 Multi-agent Reinforcement Learning in Sequential Social Dilemmas, Hughes et al. 2018 Inequity aversion improves cooperation in intertemporal social dilemmas, Du et al. 2023 A Review of Cooperation in Multi-agent Learning

How to build bridges?

Realize: focus on modeling for understanding - not scaling for engineering

All models are wrong

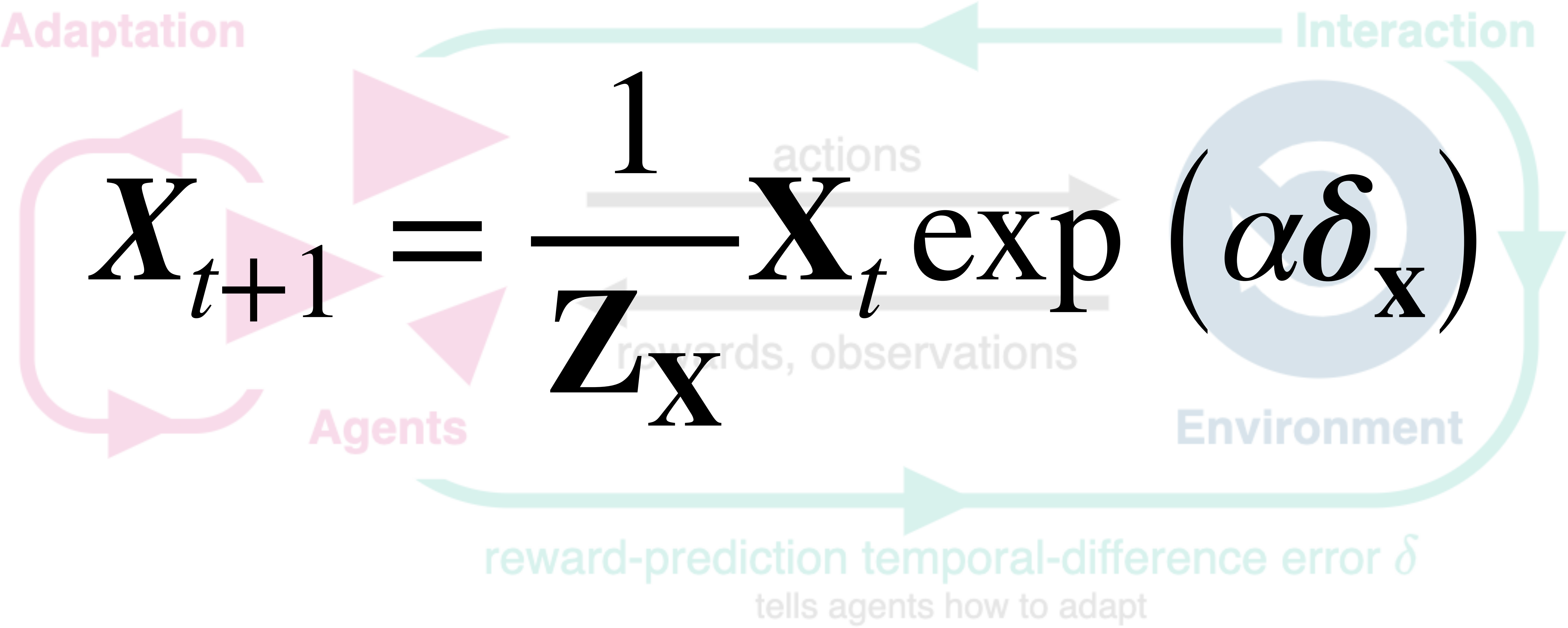
Study low-dimensional environments

Controll stochasticity

Collective reinforcement learning dynamics

Treat reinforcement learning as a nonlinear dynamical system

Deterministic approximation from strategy averaging



$$X_{t+1} = \frac{1}{Z_{\mathbf{X}}} \mathbf{X}_t \exp\left(\alpha \boldsymbol{\delta}_{\mathbf{X}}\right)$$

$$\boldsymbol{\delta} \leftarrow \boldsymbol{\delta}_{\mathbf{X}} + \epsilon$$

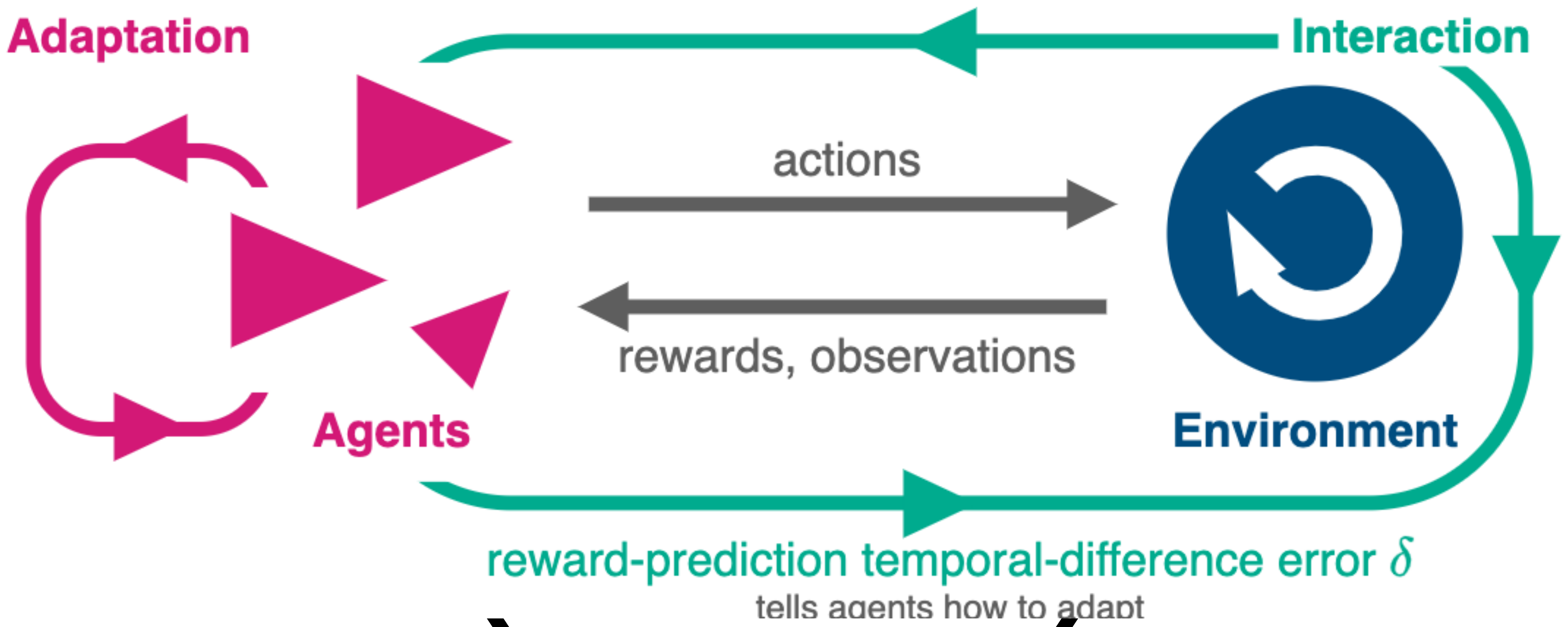
Agents learn how to act as if having a perfect model of the world

Integrating three types of models

Dynamic-system models

Equilibrium-based models

$$X_{t+1} = \frac{1}{\mathbf{Z_X}} \mathbf{X}_t \exp\left(\alpha \boldsymbol{\delta}_{\mathbf{x}}\right)$$

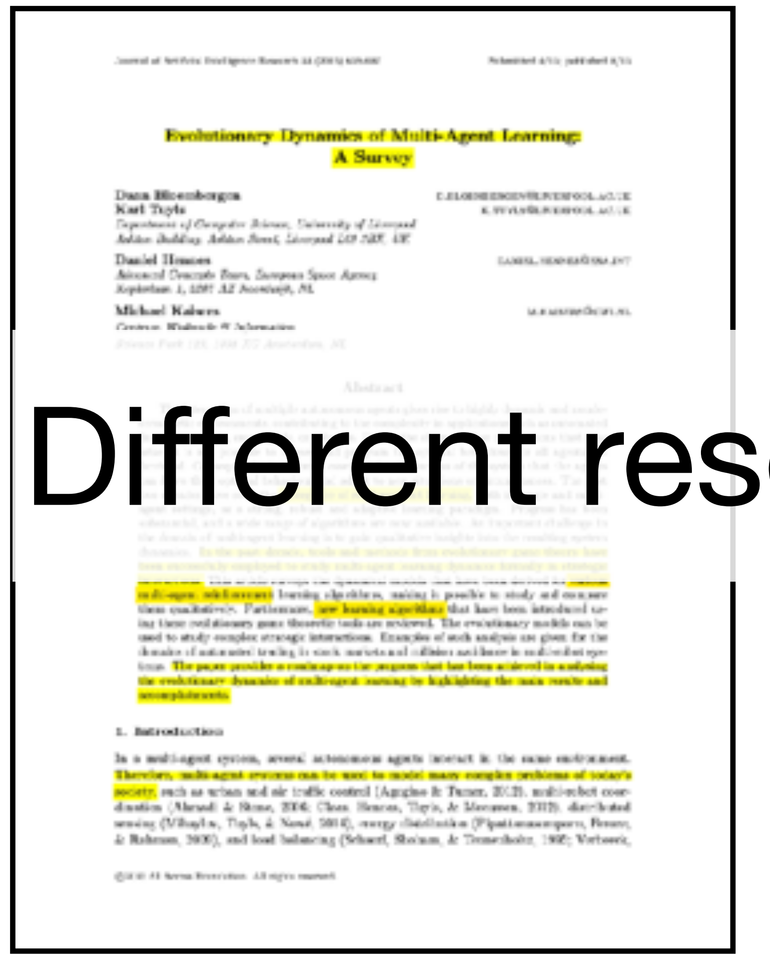


Agent-based models

Not a new idea

Building upon an interdisciplinary but scattered foundation

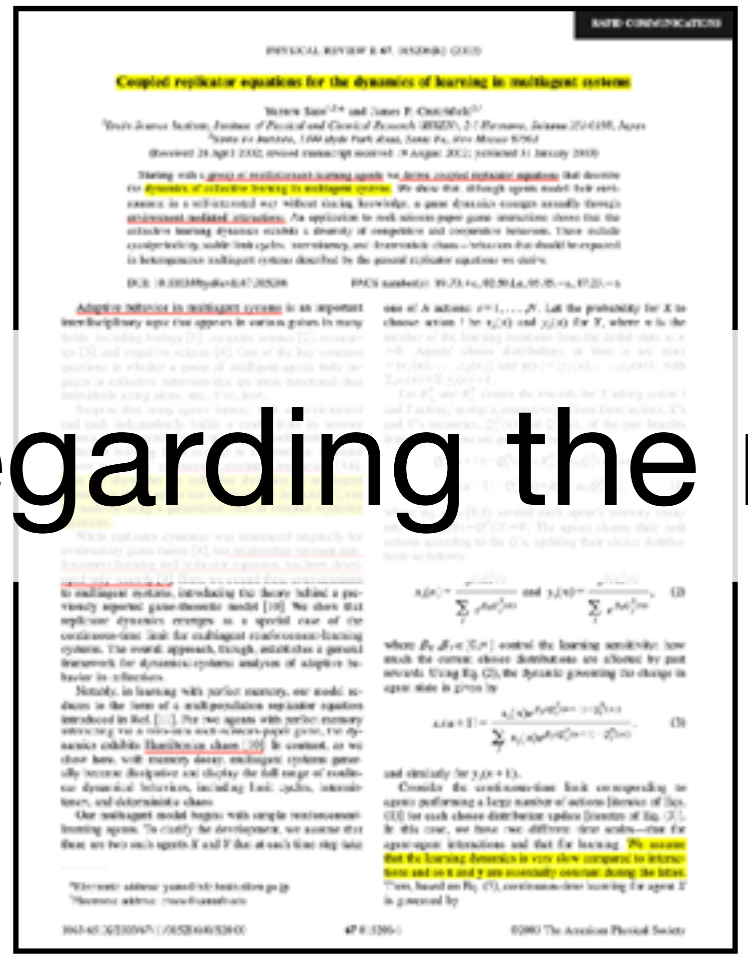
Machine Learning	Controll & Engineering	Statistical Mechanics	Economics	Sociology	Mathematical Biology


Evolutionary Dynamics of Multi-Agent Learning: A Survey

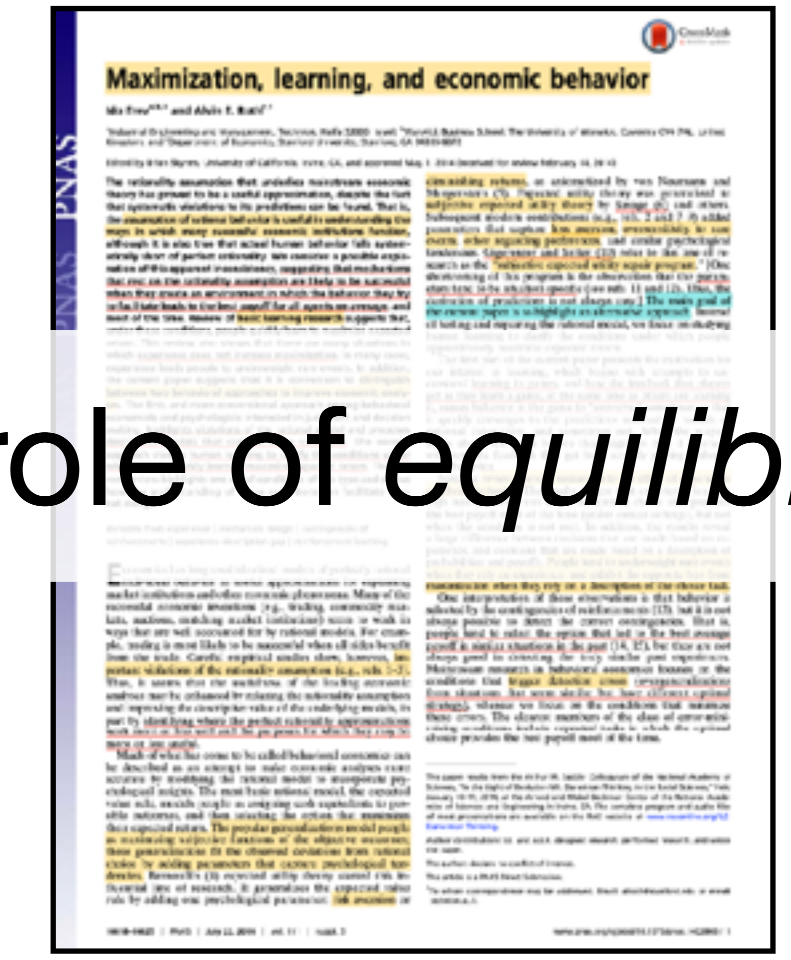
Bloembergen et al. 2015


The Confluence of Networks, Games, and Learning a Game-Theoretic Framework for Multiagent Decision Making Over Networks

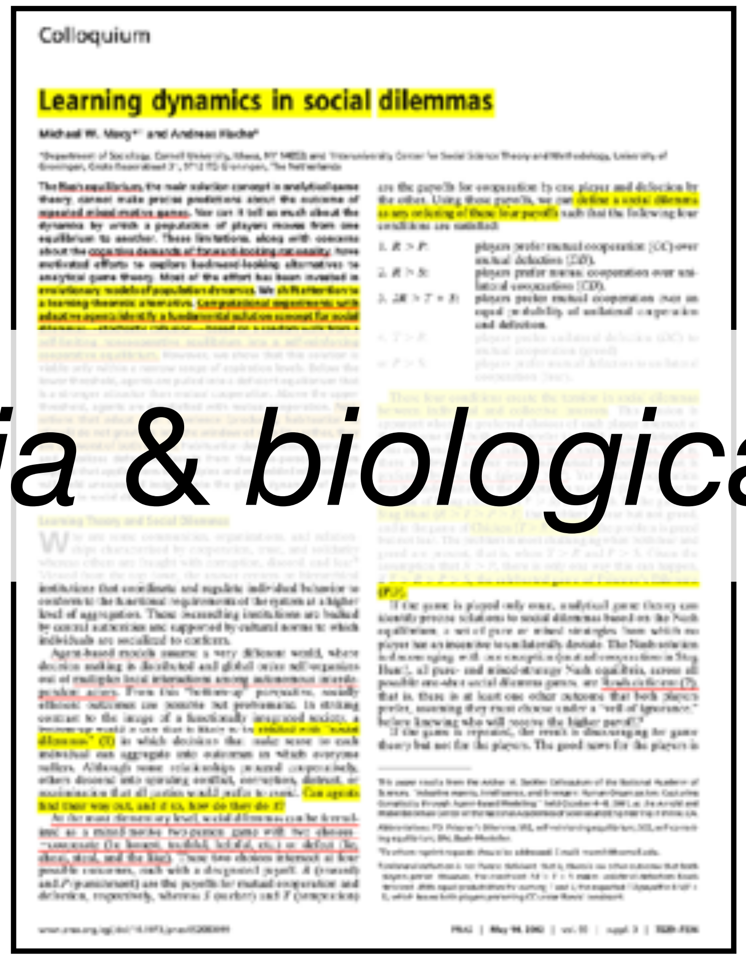
Li et al. 2022


Coupled replicator equations for the dynamics of learning in multiagent systems

Sato & Crutchfield 2003

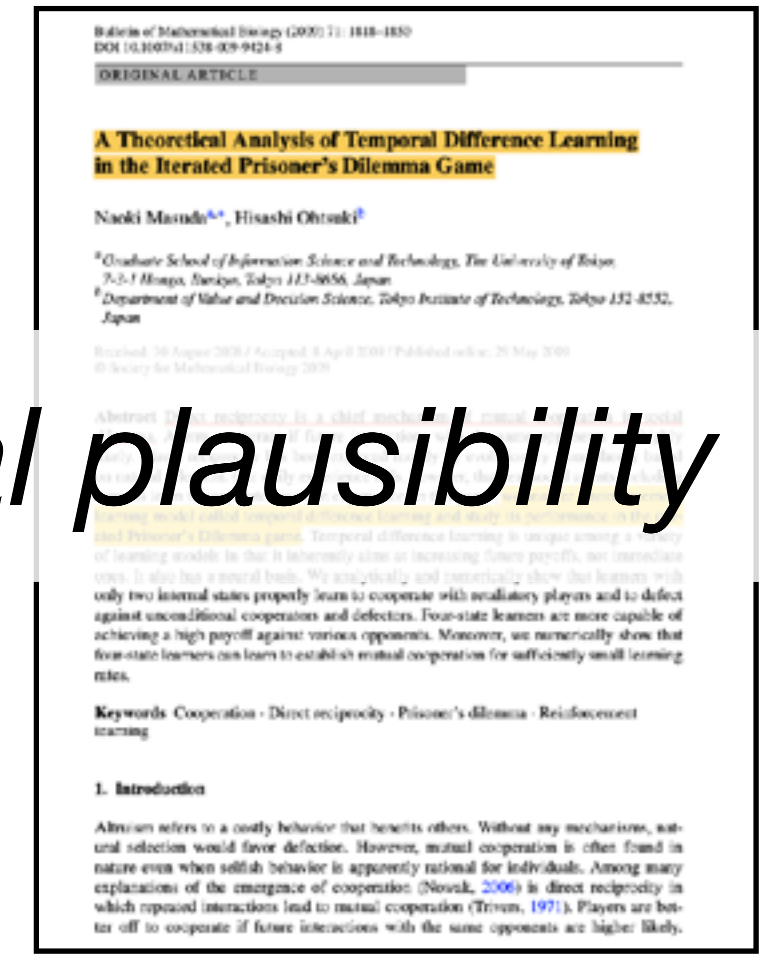

Maximization, learning, and economic behavior

Erev & Roth 2014


Colloquium

Learning dynamics in social dilemmas

Macy & Flache 2002


A Theoretical Analysis of Temporal Difference Learning in the Iterated Prisoner's Dilemma Game

Masuda & Nakamura 2011

Different research goals regarding the role of *equilibria* & *biological plausibility*

Our focus | *non-linear dynamics* of *biologically-inspired* learning of *cooperation*

WHAT | MARL→CSS
Cognition in contexts

Cooperation challenges in human-machine ecologies

Why are we not cooperating more toward a sustainable future for all?

Collective Behavior 👥
Social context
Heterogeneity
Structure
Scale

Cognitive Agency 🧠
Perception
Valuation
Decision-making
Action

Environmental context 🌳
Smooth changes
Abrupt changes
Uncertain changes
Delayed consequences

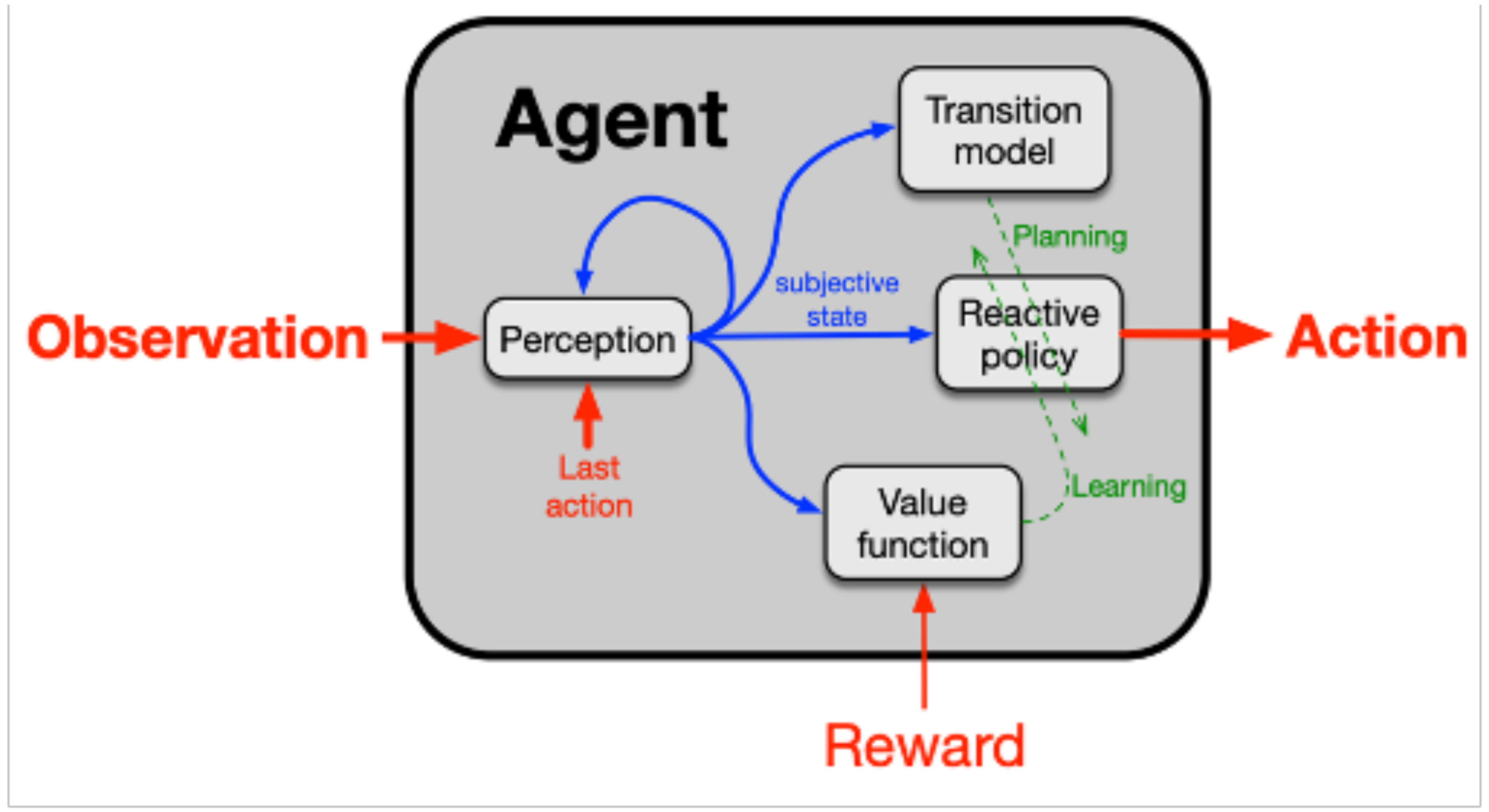
Good modeling practices 🗺️
Coherence | Transparency | Sparsity

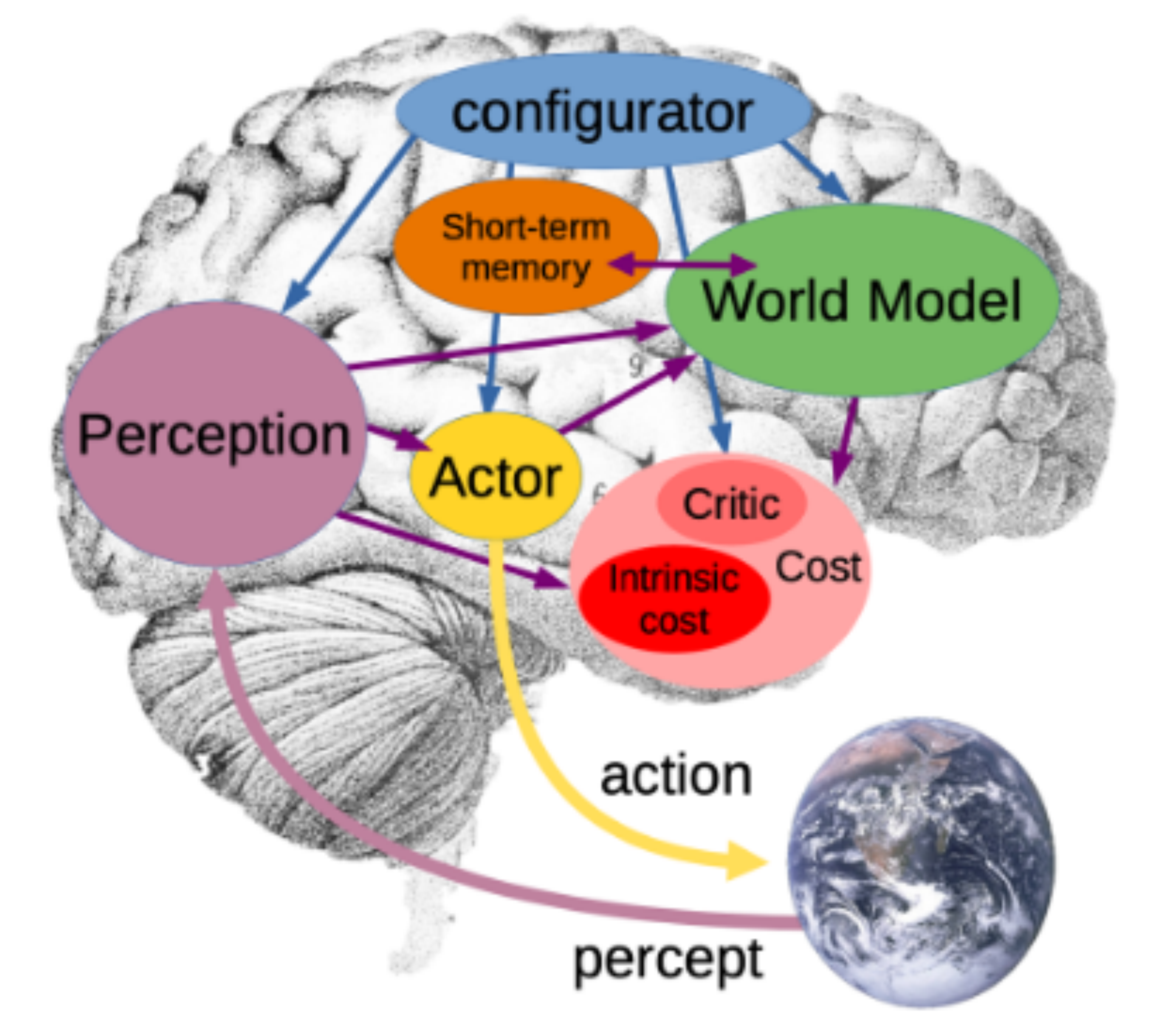
Dynamics 🚀
Nonlinearities & Feedbacks
Stability & Resilience
Transient evolution
Critical transitions

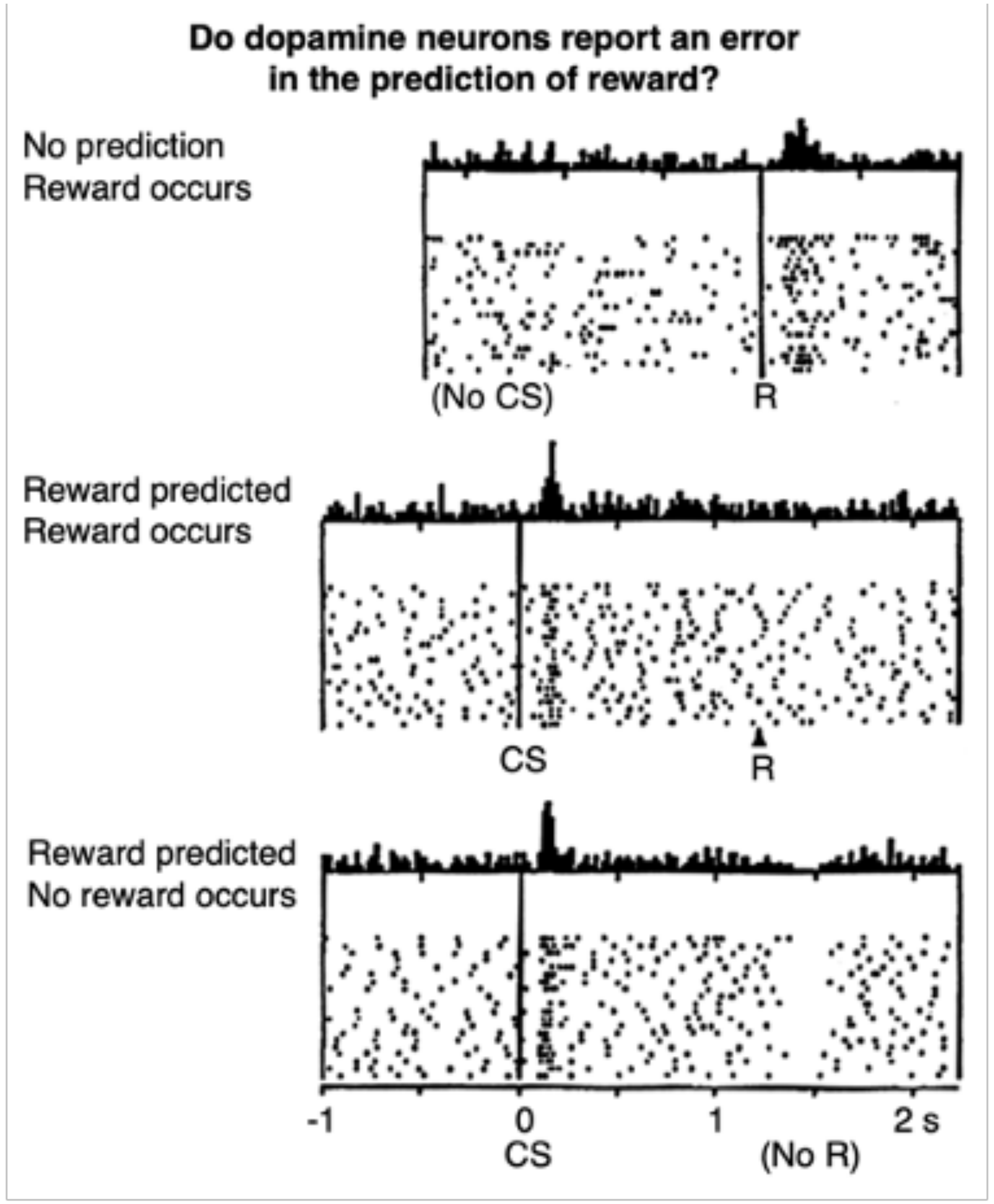
Personal synthesis based on Schill et al. (2019), Elsawah et al. (2020), Müller et al. (2020), Constantino et al. (2021), Levin & Xepapadeas (2021), Farahbakhsh et al. (2022), Giupponi et al. (2022)

Individual cognition | Intelligent and adaptive behavior

Reinforcement learning as a general prototype model for intelligent & adaptive decision-making



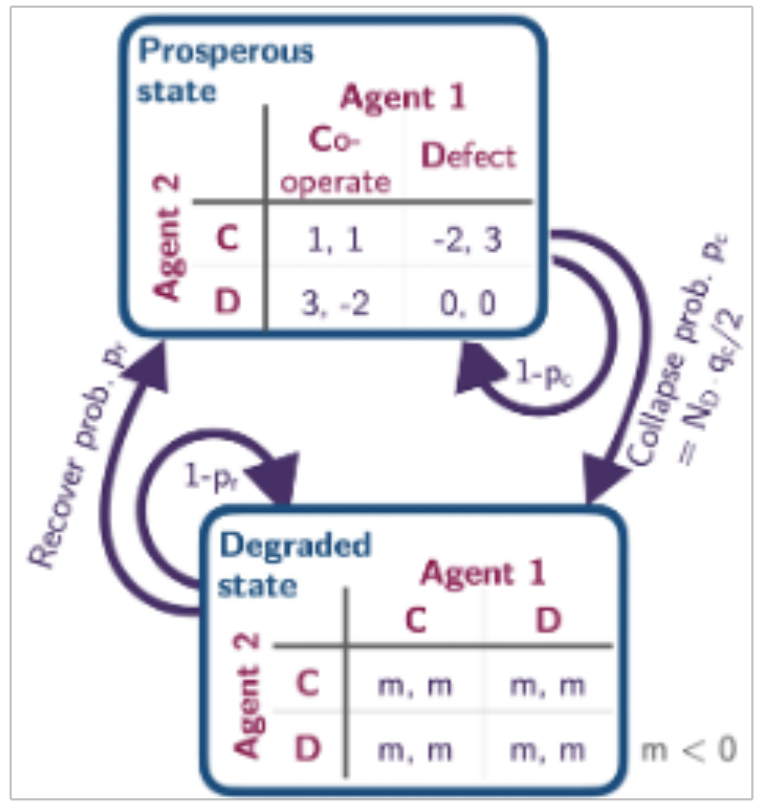




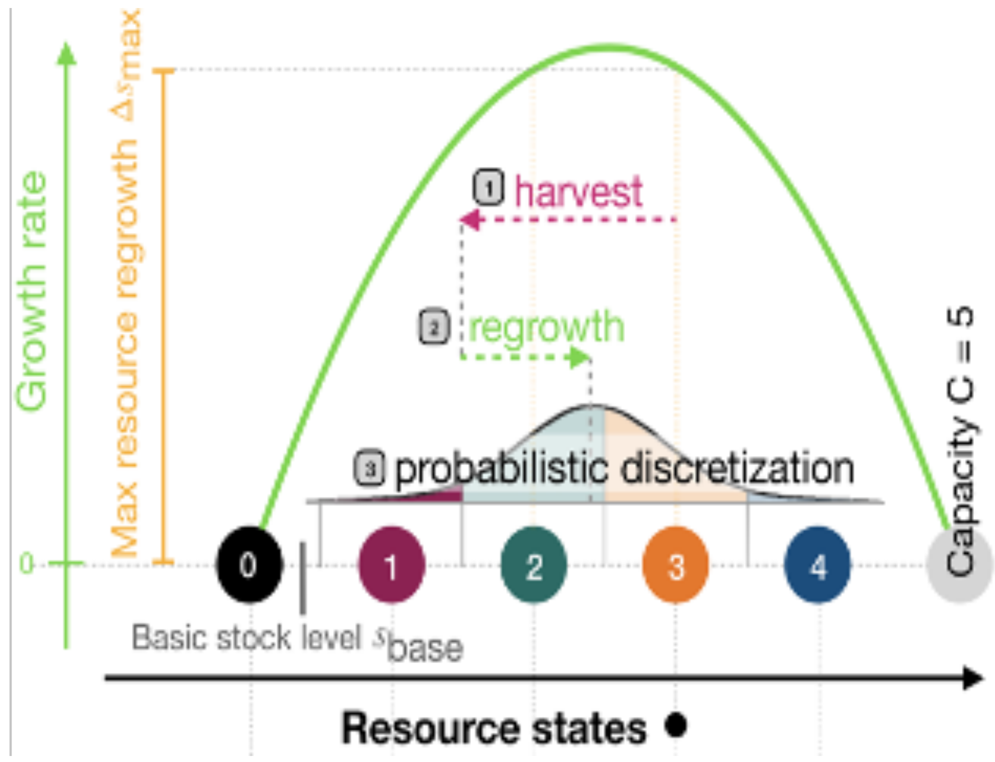
Temporal-difference reinforcement learning empirically grounded in biology

Embedded cognition | Ecology & environment

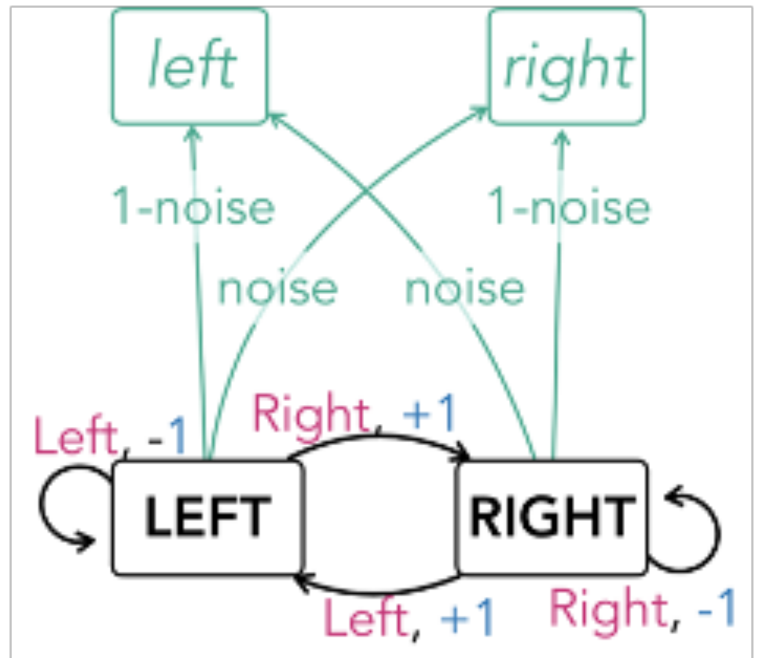
Partially observable stochastic games as a general model for environmental context with delayed and stochastic consequences



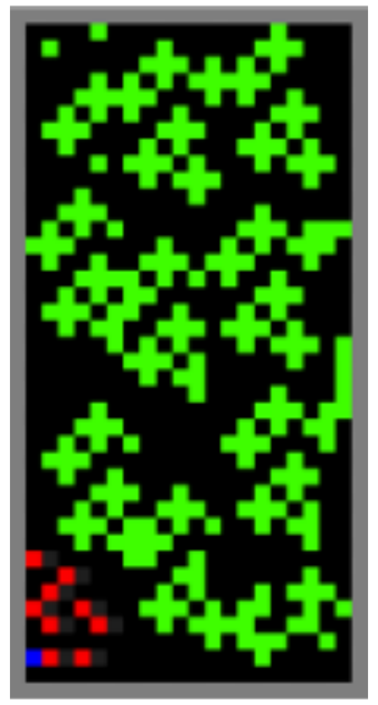
Abrupt transitions



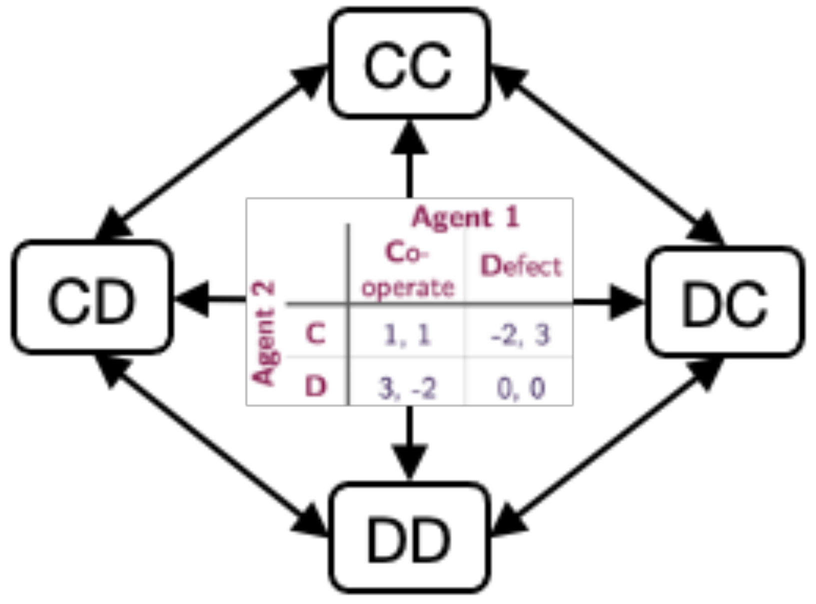
Smooth dynamics



Partial observability



Spatial



Social environment

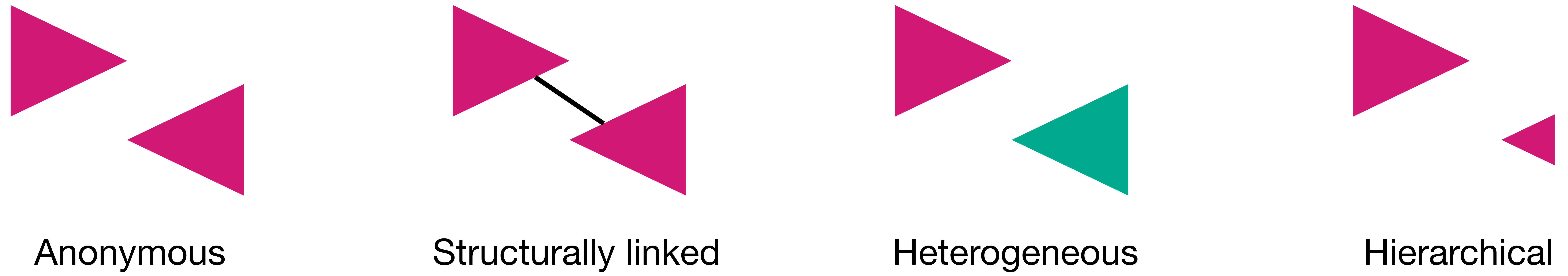
Pure strategic interactions or noisy reward feedback included

Discretization powerful - numerically and conceptually

Barfuss et al. 2020 Caring for the future can turn tragedy into comedy for long-term collective action under risk of collapse, Barfuss & Mann 2022 Modeling the effects of environmental and perceptual uncertainty using deterministic reinforcement learning dynamics with partial observability, Perolat et al. 2017 A multi-agent reinforcement learning model of common-pool resource appropriation, Barfuss & Meylahn 2023 Intrinsic fluctuations of reinforcement learning promote cooperation

Collective cognition | Multi-agent systems

Multi-agent systems as a general model to micro-found collective behavior



Large collectives cause a curse of dimensionality

Mean-field approaches can help

WHAT | CSS→MARL
Emergent phenomena

Cooperation challenges in human-machine ecologies

Why are we not cooperating more toward a sustainable future for all?

Collective Behavior 👥
Social context
Heterogeneity
Structure
Scale

Cognitive Agency 🧠
Perception
Valuation
Decision-making
Action

Good modeling practices 🗺️
Coherence | Transparency | Sparsity

Environmental context 🌳
Smooth changes
Abrupt changes
Uncertain changes
Delayed consequences

Dynamics 🚀
Nonlinearities & Feedbacks
Stability & Resilience
Transient evolution
Critical transitions

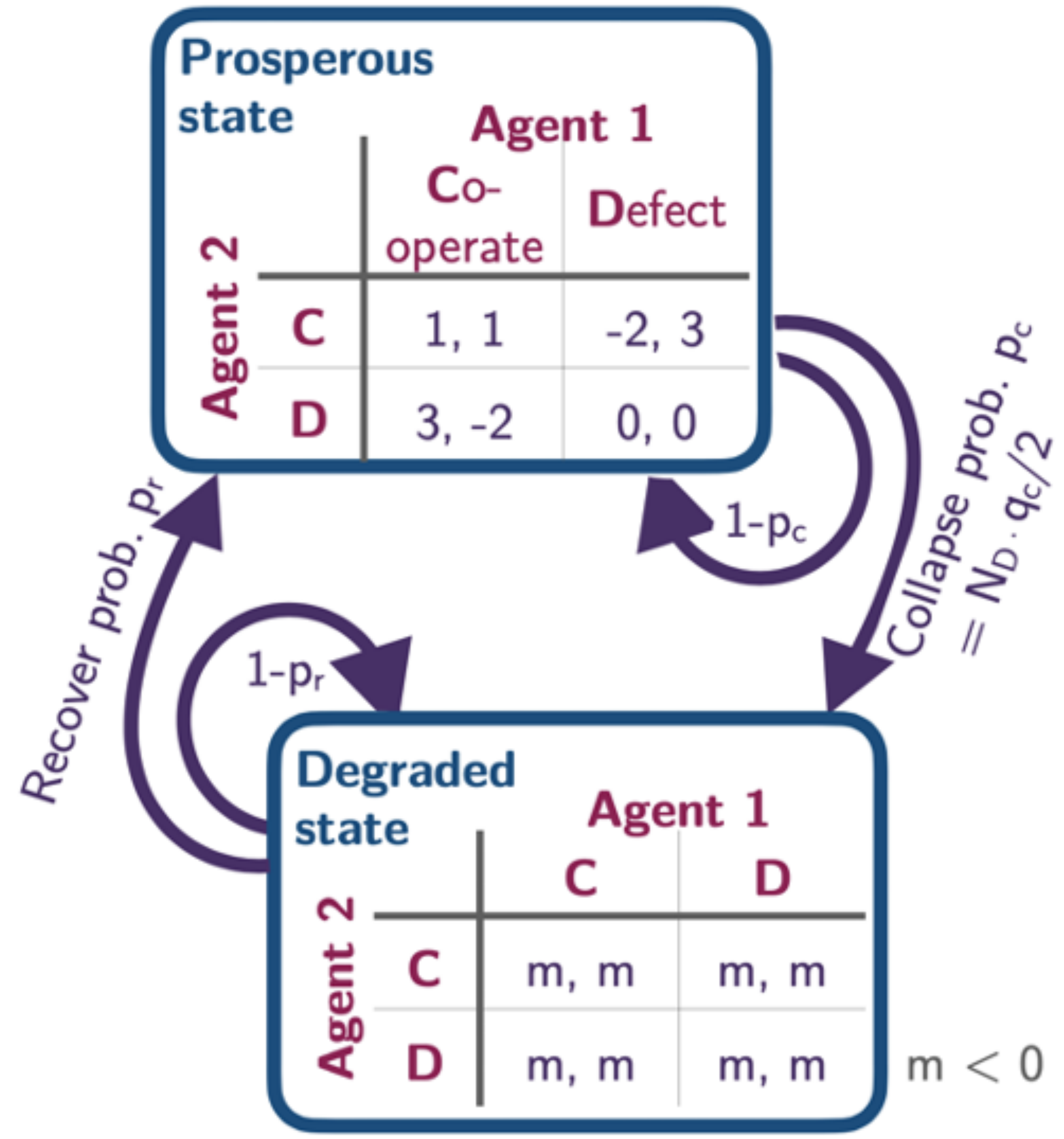
Personal synthesis based on Schill et al. (2019), Elsawah et al. (2020), Müller et al. (2020), Constantino et al. (2021), Levin & Xepapadeas (2021), Farahbakhsh et al. (2022), Giupponi et al. (2022)

A minimal model of

Collective action

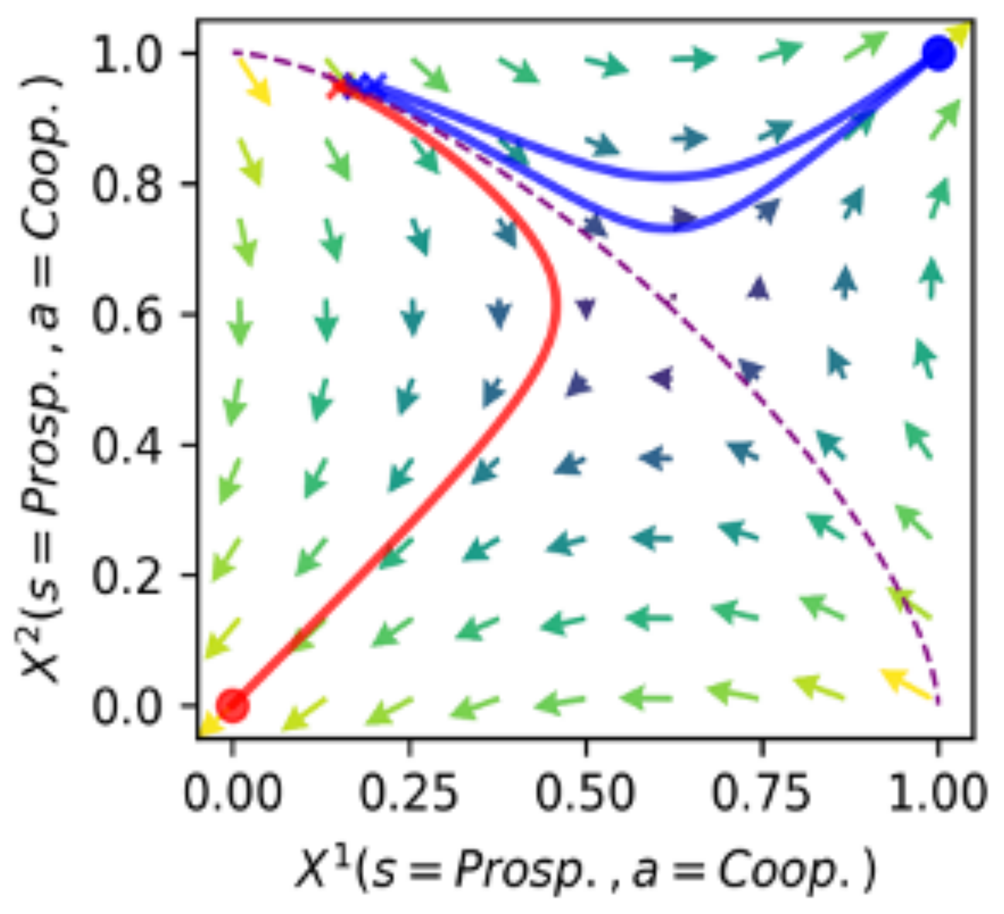
from **individual** adaptive intelligent decision-making via RL

in **environmental** commons with catastrophic thresholds

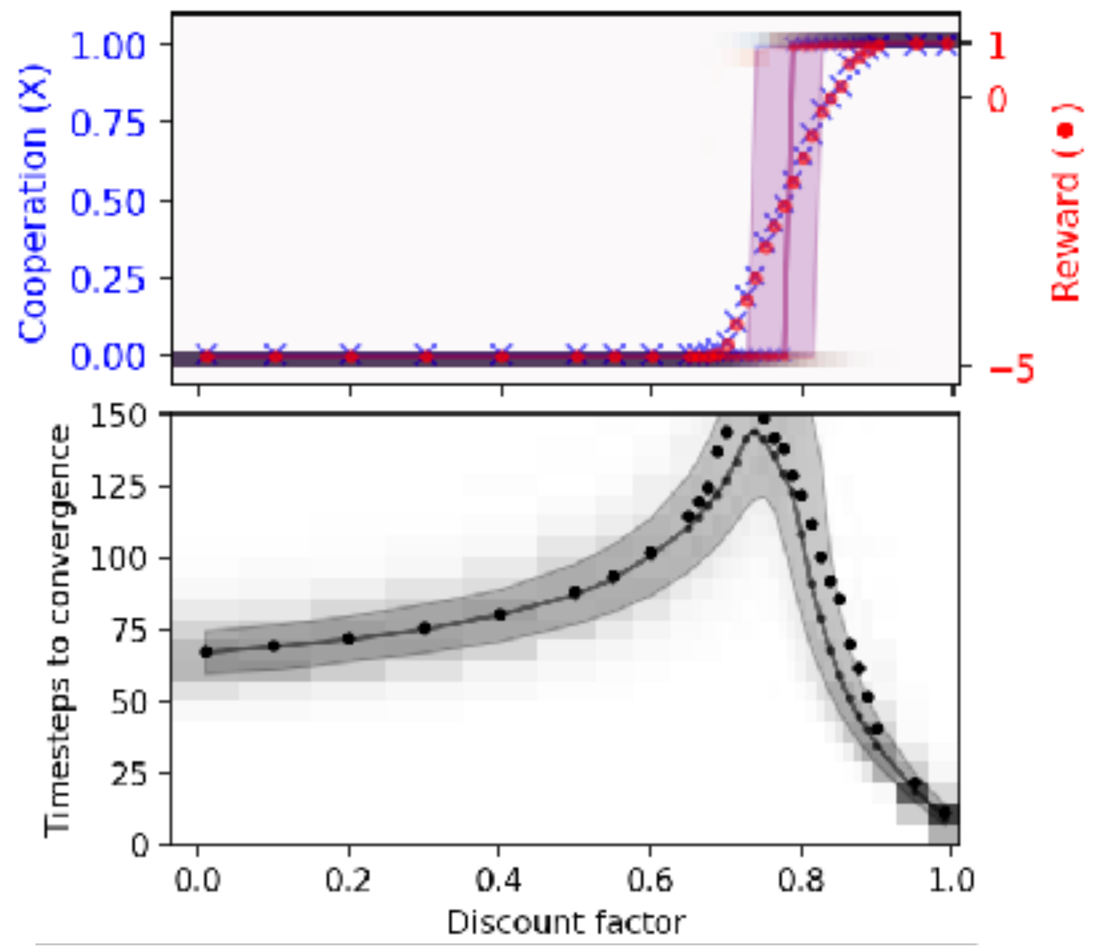


Barfuss et al. 2020 Caring for the future can turn tragedy into comedy for long-term collective action under risk of collapse

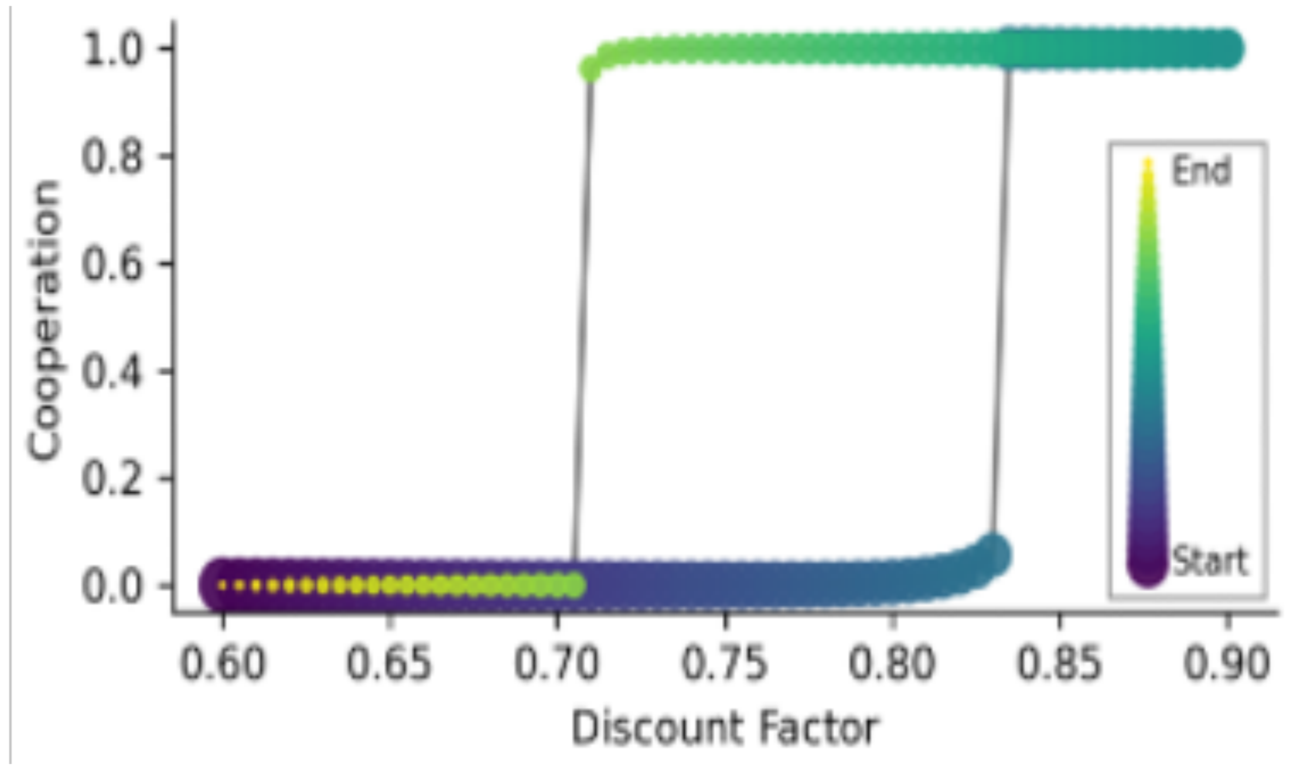
Complex emergent phenomena



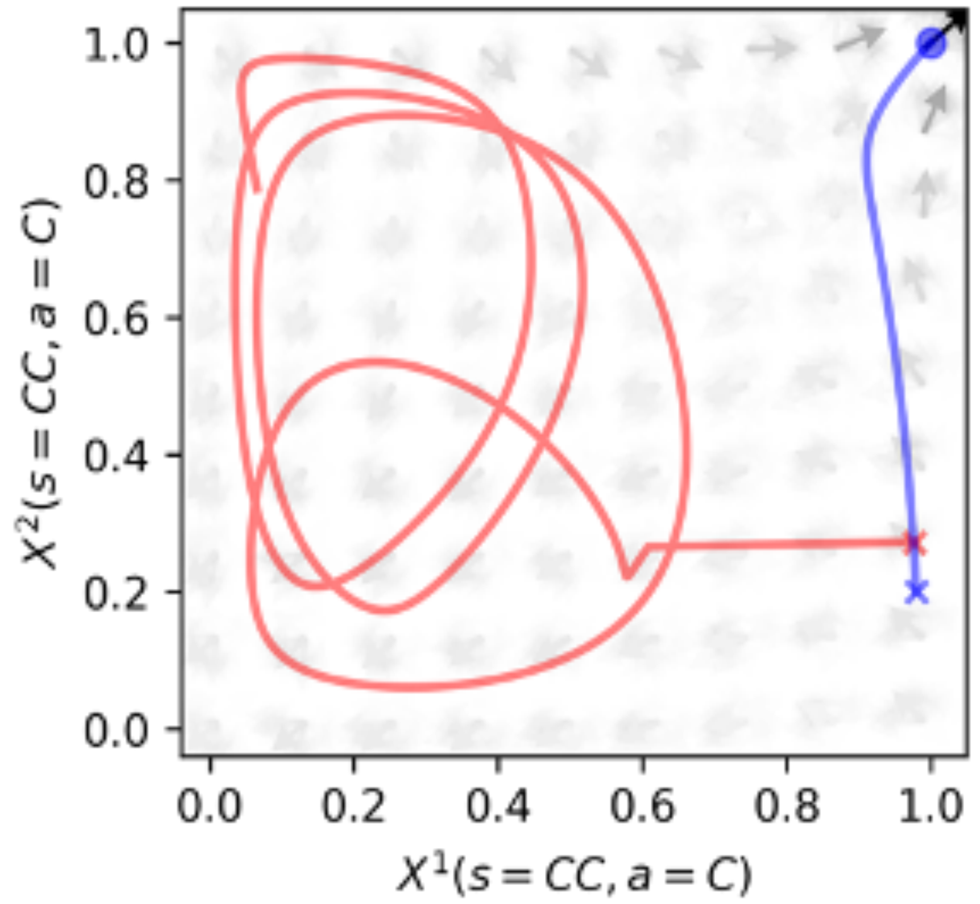
Multi-Stability



Abrupt Transitions



Hysteresis



Dynamic Regimes

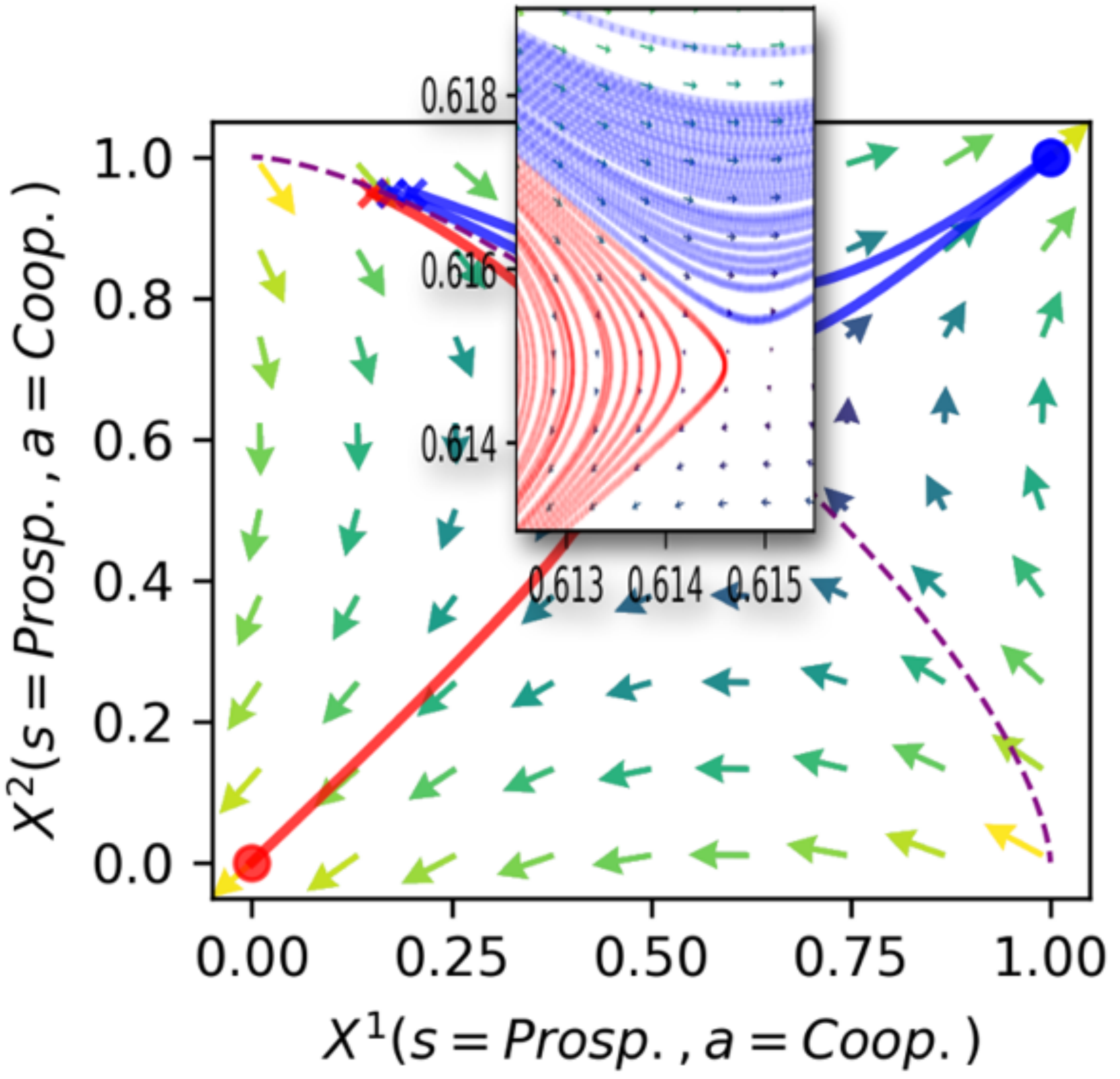
Multi-stability in commons with catastrophic thresholds

Basin of attraction (of the cooperative equilibrium) as a measure of **collective intelligence** Leonard & Levin, 2022

Encoding information in equilibrium strategy as **emergent collective memory**

Geometric view for **ad-hoc teamwork**

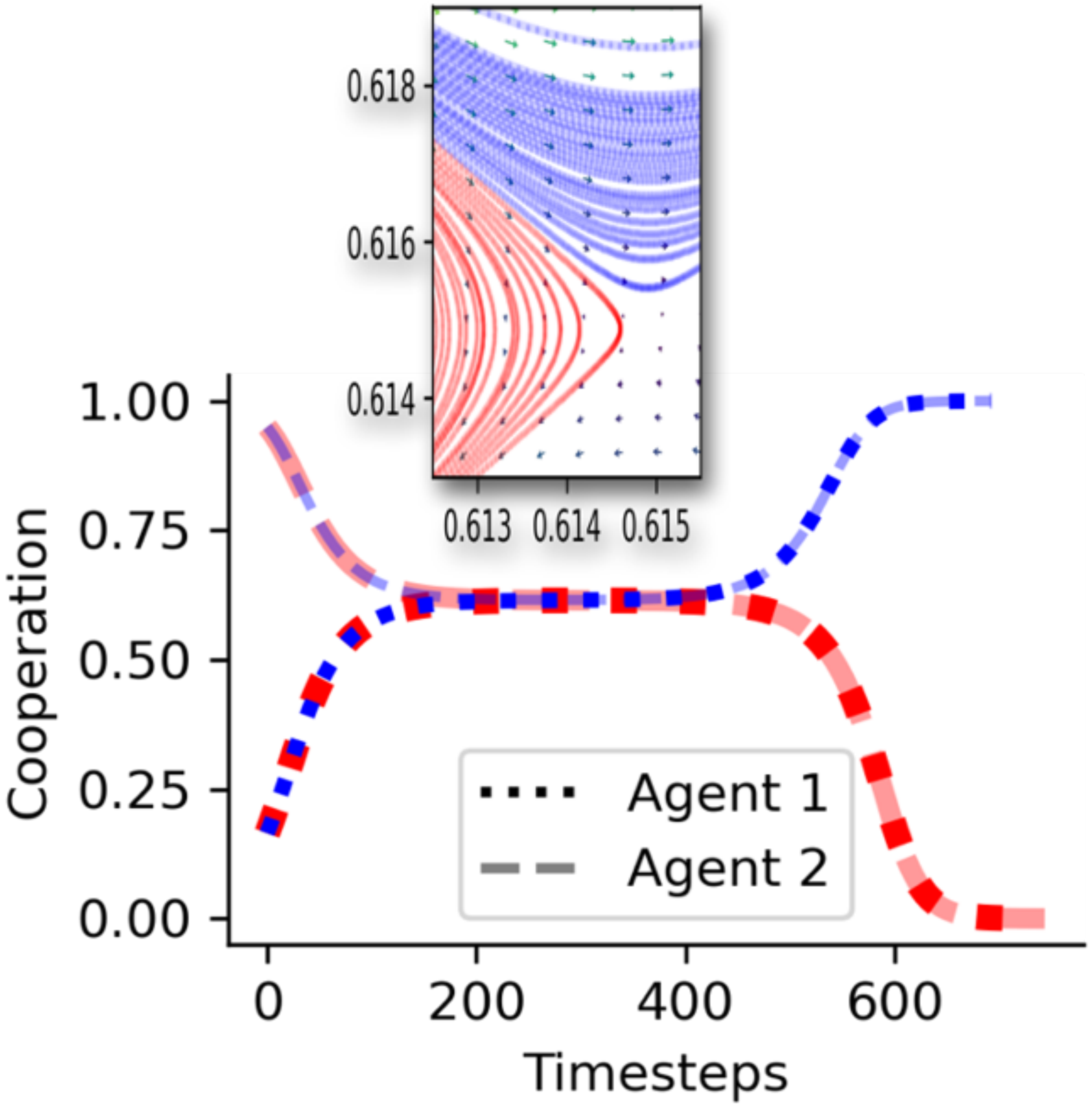
Basin of attraction as a measure of (social-ecological) **resilience**



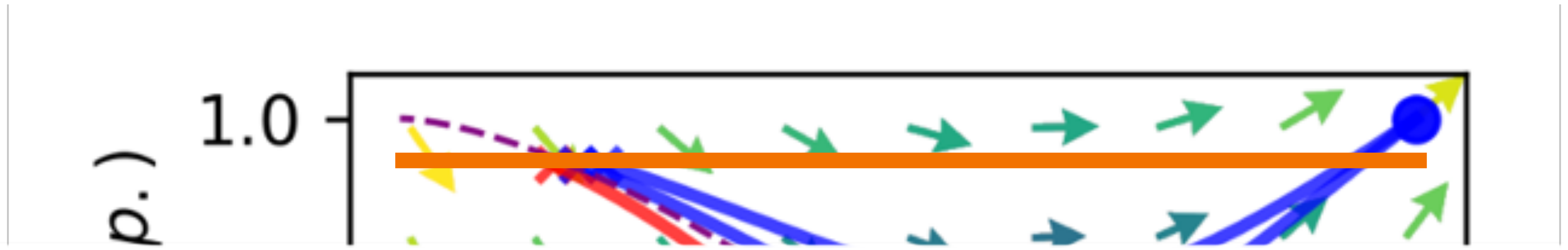
Multi-stability in commons with catastrophic thresholds

Agents' learning almost converged until suddenly drastic and fast changes may happen again

Emergent separation of time scales

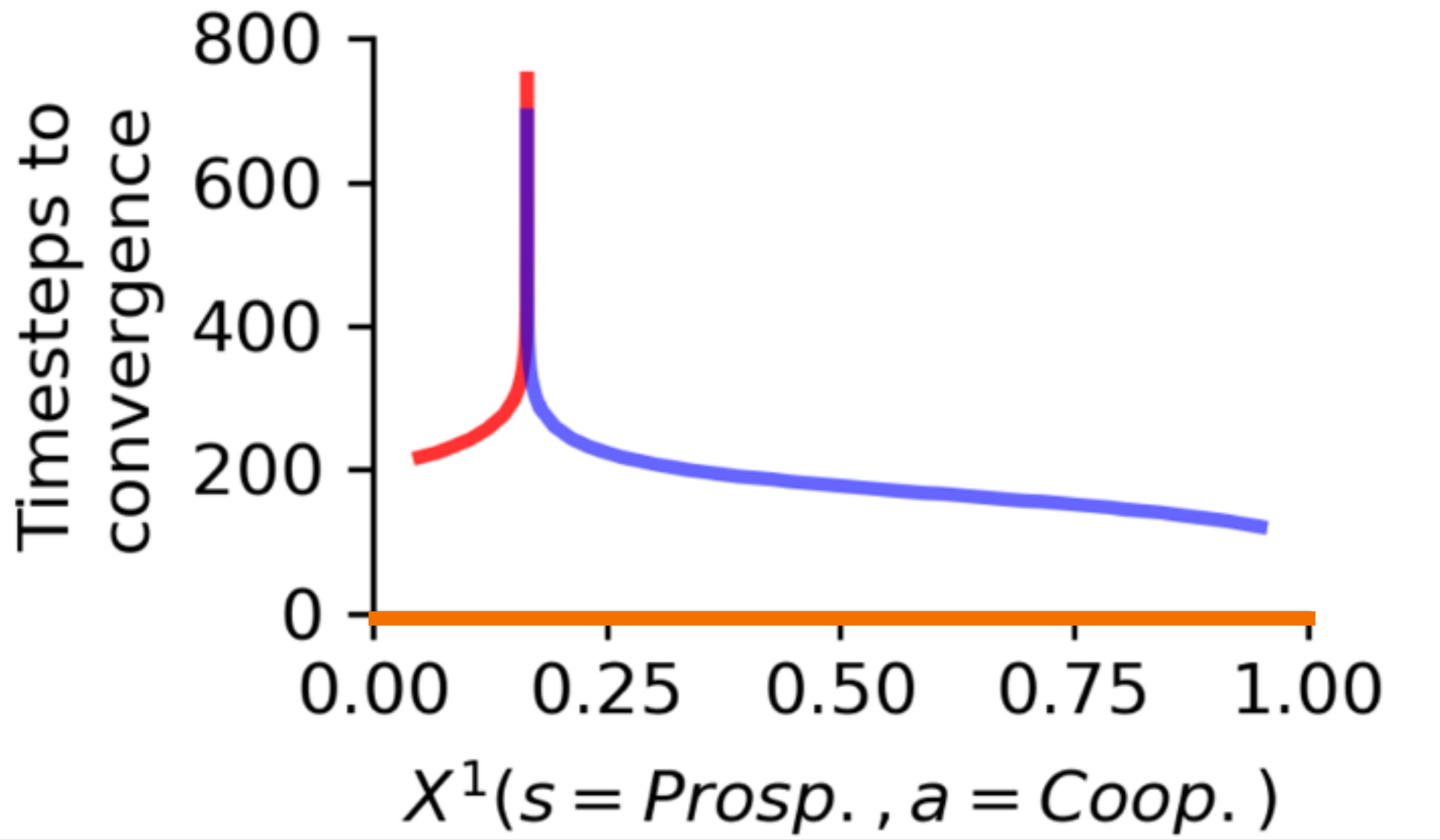


Multi-stability in commons with catastrophic thresholds



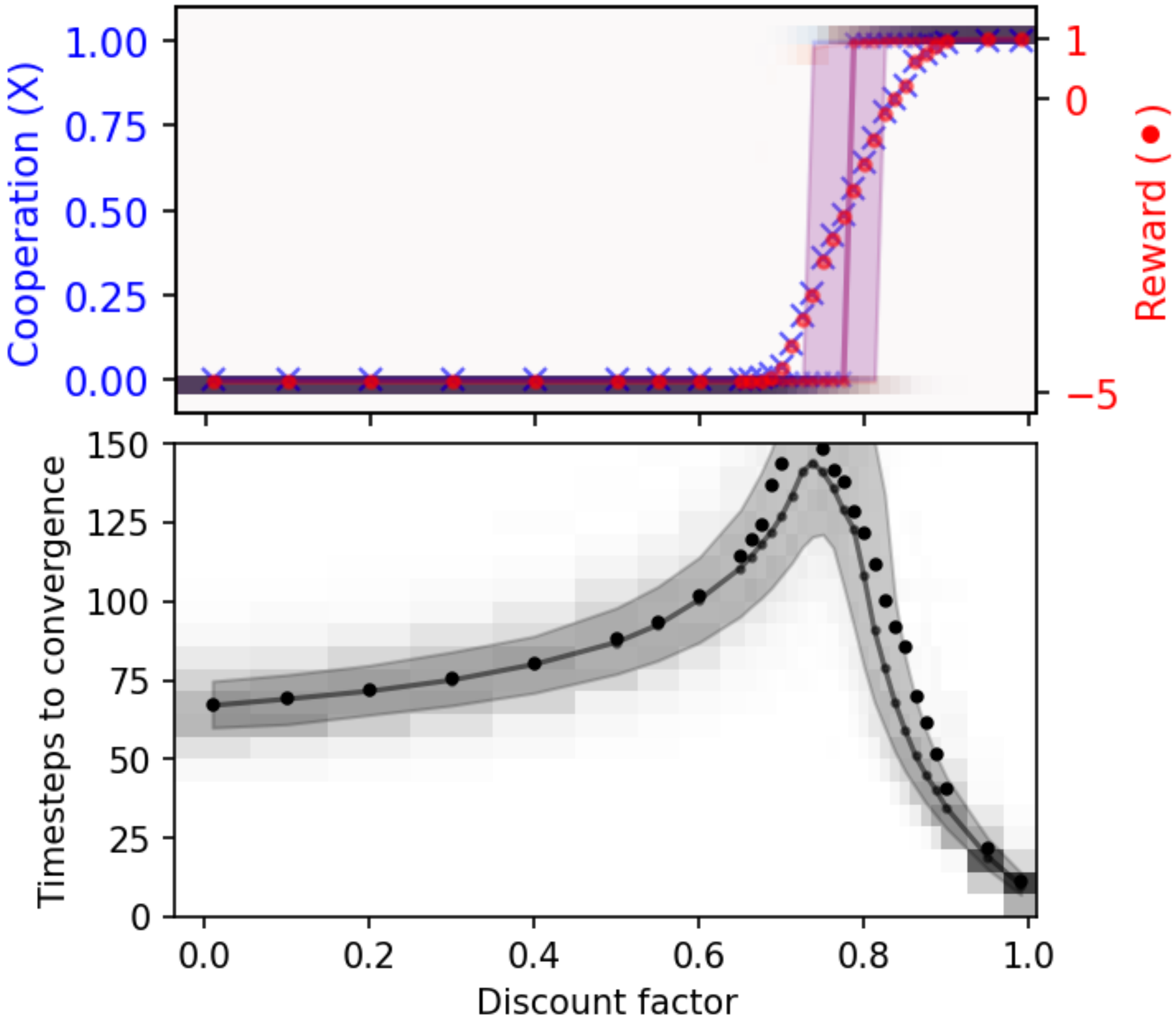
Critical slowing down at the social tipping point

- convergence less relevant
- transient learning dynamics
- early warning indicators



Critical transitions in commons with thresholds

when agents become more future-caring



- convergence less relevant
- transient learning dynamics
- early warning indicators

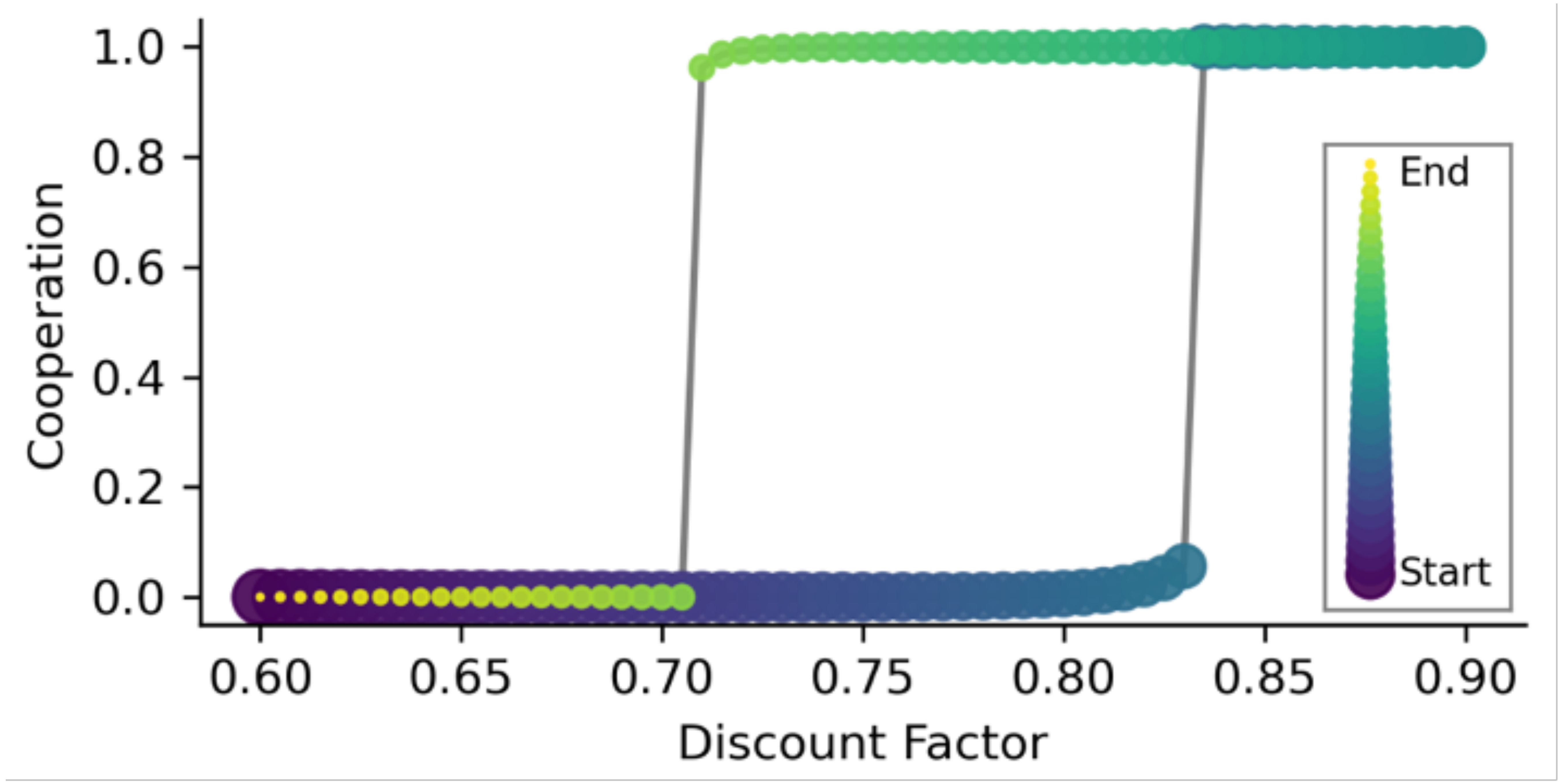
Hysteresis

e.g., when agent's future-caring weakens again

Another form of collective memory

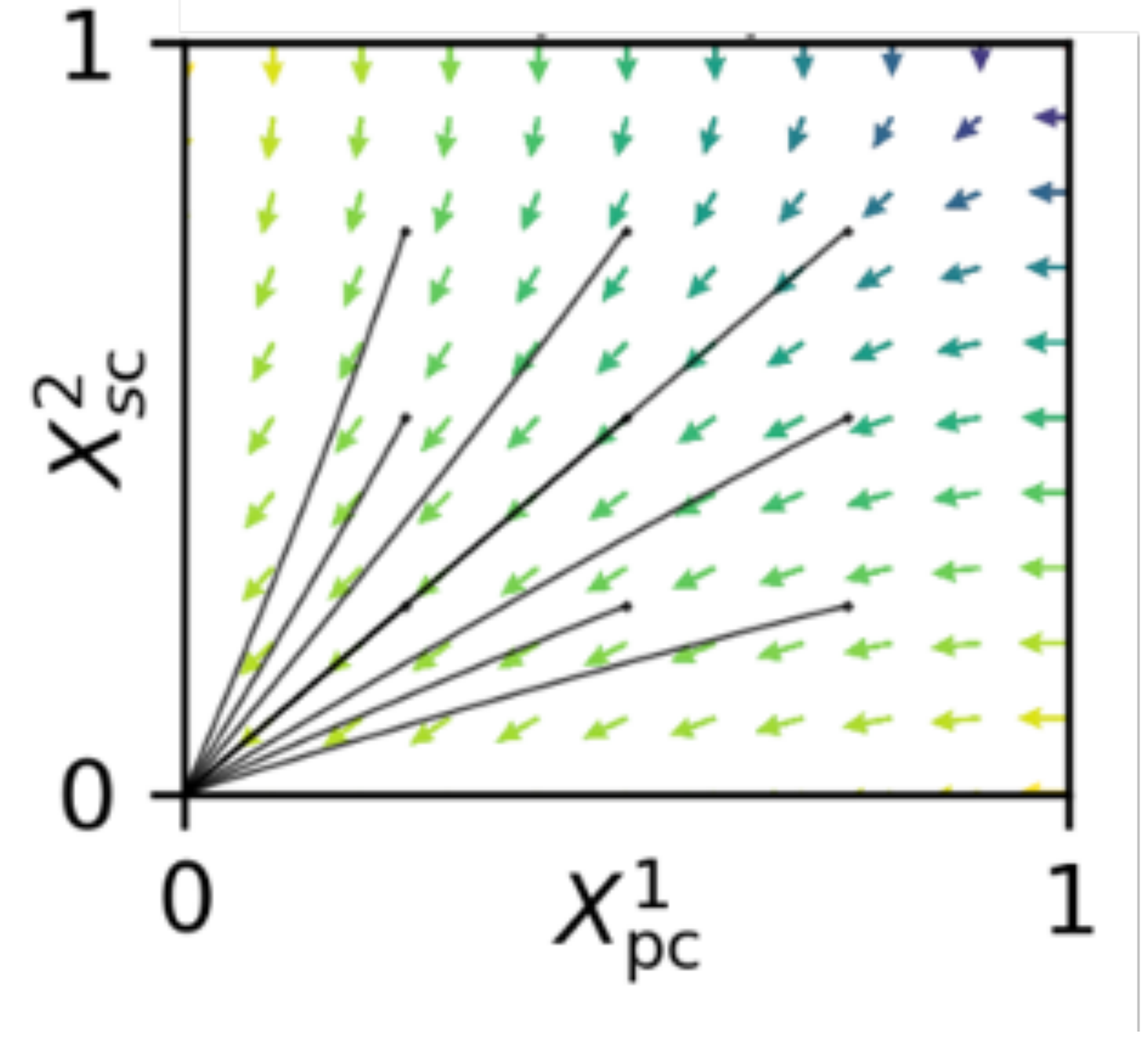
Important for social tipping elements

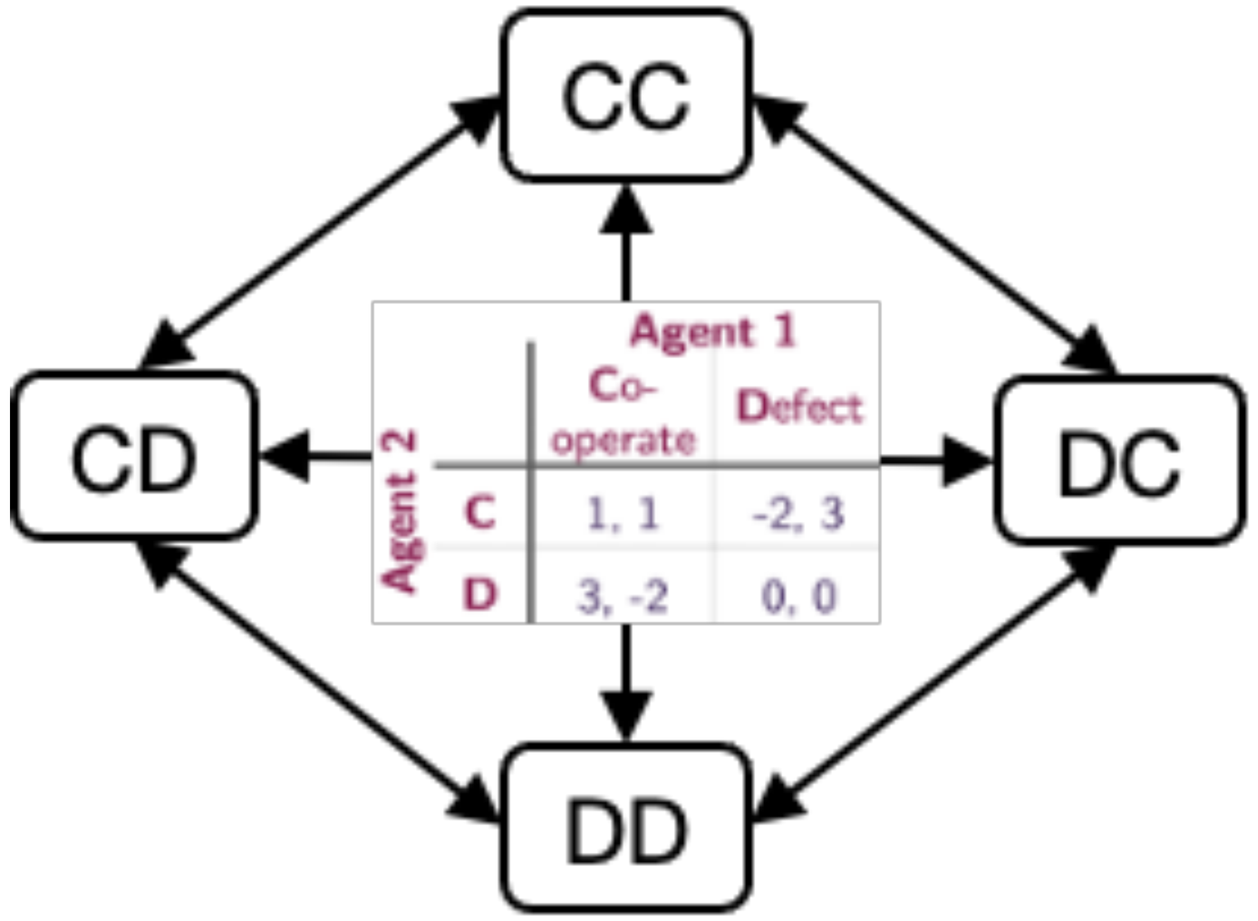
Potential for sustainability interventions

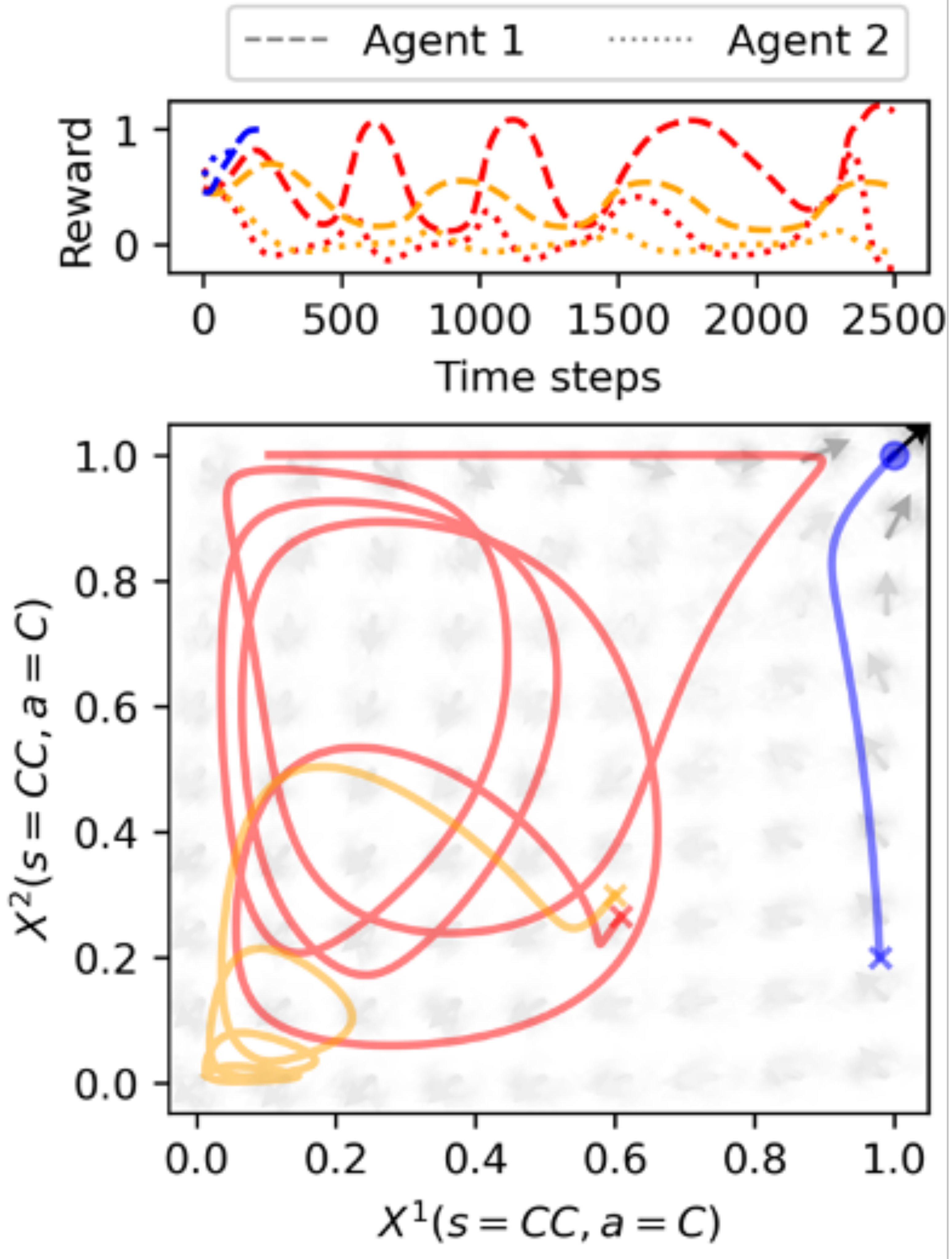


Complex collective action dynamics in social dilemma

Agent 2 \ Agent 1	Co-operate	Defect
C	1, 1	-2, 3
D	3, -2	0, 0







When agents condition on the previous round's action, they can

- learn to cooperate
- learn on oscillating and unpredictable transients

CONCLUSION

PNAS PERSPECTIVE in press OPEN ACCESS

Check for updates

Collective cooperative intelligence

Wolfram Barfuss[a,1], Jessica Flack[b], Chaitanya S. Gokhale[c,d], Lewis Hammond[e], Christian Hilbe[c], Edward Hughes[f], Joel Z. Leibo[f], Tom Lenaerts[g,h,i], Naomi Leonard[j], Simon Levin[k], Udari Madhushani Sehwag[l,2], Alex McAvoy[m,n], Janusz M. Meylahn[o], and Fernando P. Santos[p]

Edited by Marco A. Janssen, Arizona State University, Tempe, AZ; received December 12, 2023; accepted May 6, 2024 by Editorial Board Member Elke U. Weber

Cooperation at scale is critical for achieving a sustainable future for humanity. However, achieving collective, cooperative behavior—in which intelligent actors in complex environments jointly improve their well-being—remains poorly understood. Complex systems science (CSS) provides a rich understanding of collective phenomena, the evolution of cooperation, and the institutions that can sustain both. Yet, much of the theory in this area fails to fully consider individual-level complexity and environmental context—largely for the sake of tractability and because it has not been clear how to do so rigorously. These elements are well captured in multiagent reinforcement learning (MARL), which has recently put focus on cooperative (artificial) intelligence. However, typical

or reward scheme, e.g., via taxes and subsidies, that makes selfish actions less attractive to individuals, whereas bottom–up arrangements and social reciprocity find a way to punish defecting behavior through peers (4). However, the challenge of cooperation is far from being solved.

First, large collectives complicate the emergence and robustness of cooperation. Although many mechanisms have been identified that support its emergence and maintenance, it is also widely recognized that effective scaling mechanisms are rare (5): in global public goods, such as the climate, there is no single outside actor with sufficient enforcement power to ensure cooperation authoritatively. In situations involving many, mostly anonymous, participants, reciprocity mechanisms are hard to stabilize (6). Hence, a key

pyCRLD work-in-progress Python package

Collective Reinforcement Learning Dynamics (CRLD)

https://barfusslab.github.io/pyCRLD/

pyCRLD

CRLD

Agents
- Strategy Actor-Critic
- Strategy SARSA
- Strategy AC (part. Obs.)
- Value SARSA
- Strategy Base
- Value Base
- Strategy Base (part. Obs.)
- Base (part. Obs.)
- Base

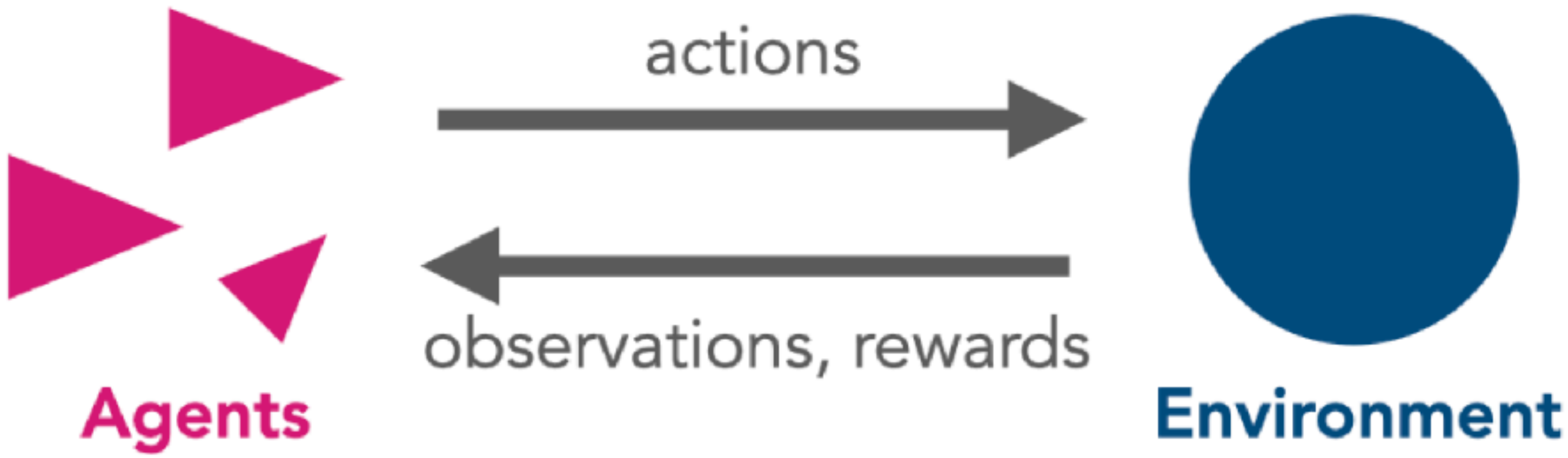
Environments
- Environment Base
- History Embedding
- Social Dilemma
- Ecological Public Good
- Uncertain Social Dilemma

CRLD

Collective Reinforcement Learning Dynamics in Python

is a tool to model the collective dynamics emerging from multi-agent reinforcement learning.

Multi-agent reinforcement learning (MARL) provides a comprehensive framework for studying the interplay among learning, representation, and decision-making between multiple actors. As a result, it offers an integrating platform to in-silico test hypotheses and build theory on how different cognitive mechanisms affect collective adaptive behavior in complex environments.

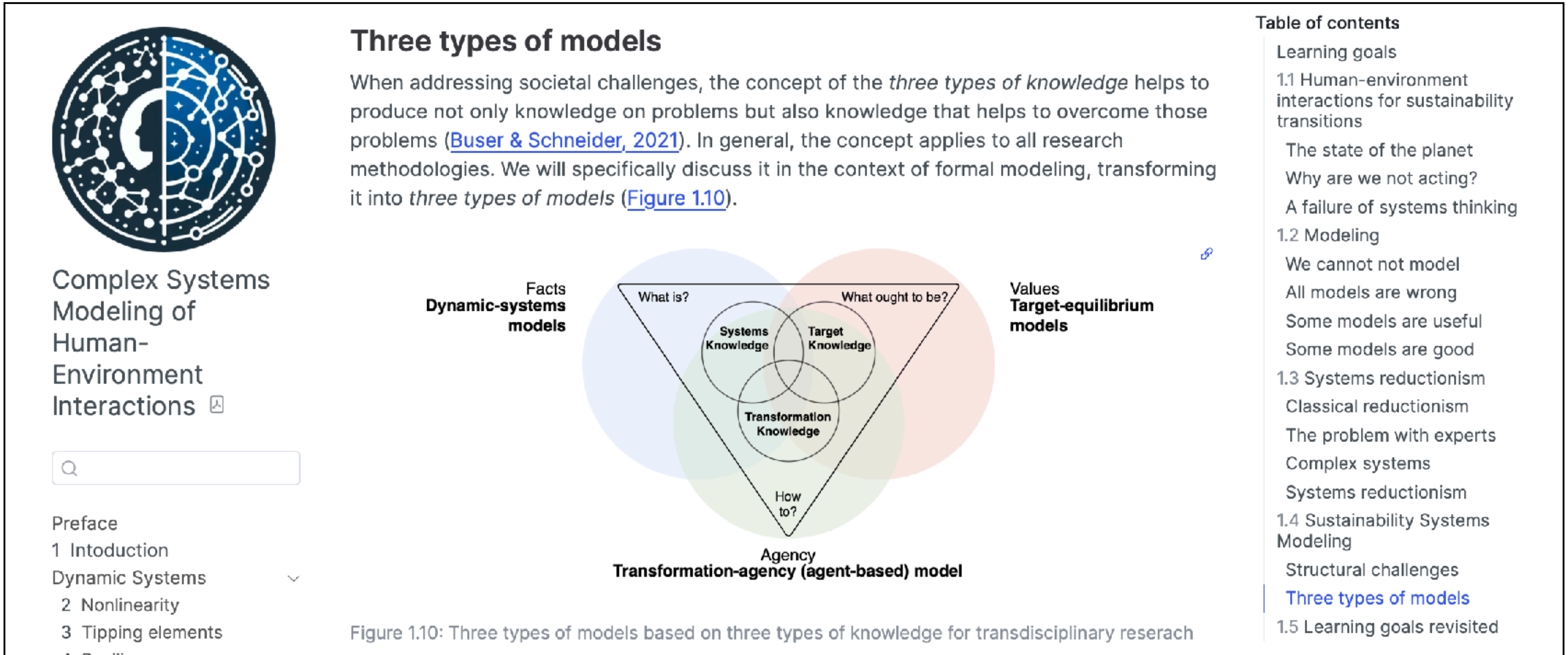


On this page
- Install
- How to use

CSMofHEI course material

Complex Systems Modeling of Human-Environment Interactions

https://wbarfuss.github.io/csm-of-hei/



Complex Systems Modeling of Human-Environment Interactions

Preface
1 Intoduction
Dynamic Systems
2 Nonlinearity
3 Tipping elements

Three types of models

When addressing societal challenges, the concept of the *three types of knowledge* helps to produce not only knowledge on problems but also knowledge that helps to overcome those problems (Buser & Schneider, 2021). In general, the concept applies to all research methodologies. We will specifically discuss it in the context of formal modeling, transforming it into *three types of models* (Figure 1.10).

Figure 1.10: Three types of models based on three types of knowledge for transdisciplinary reserach

Table of contents

Summary | CSS+MARL ⟶ SusEcon

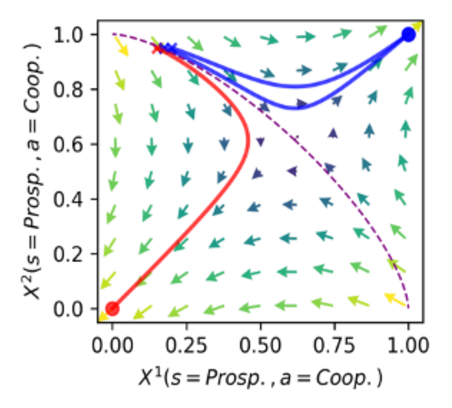
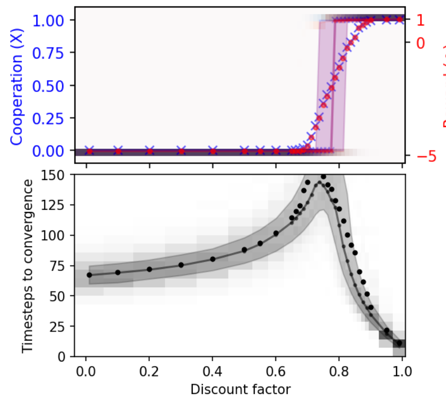
Why | Collective action challenges in human-machine ecologies

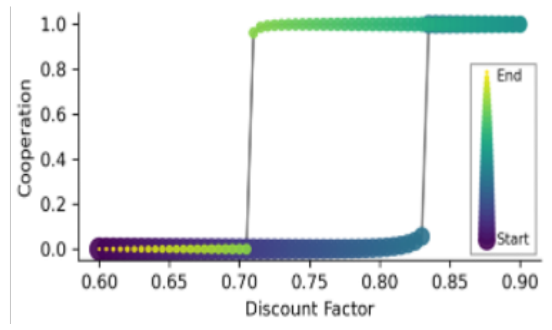
- Why are we not cooperating towards a sustainable future?
- Transparent Analysis 🔍 of the Non-linear dynamics 🚀 of the Collective behavior 👥 emerging from Individual intelligence 🧠 in some Environmental context 🌳

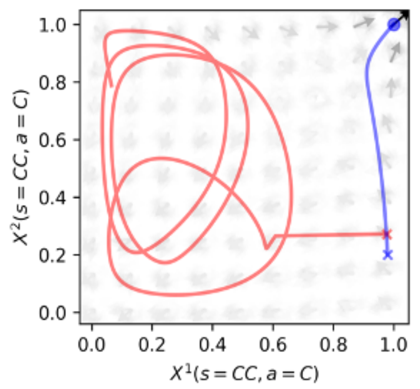
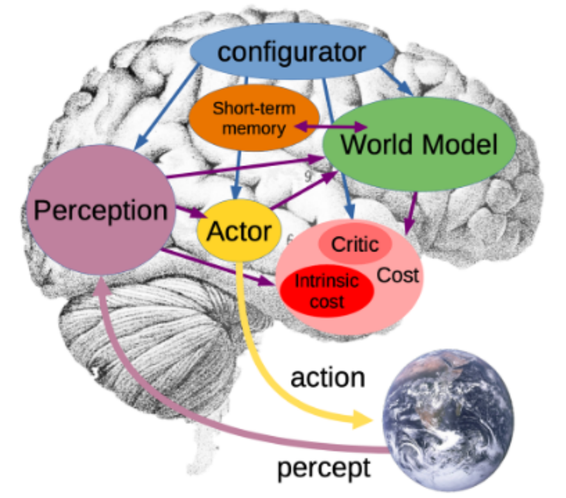
How | Building bridges

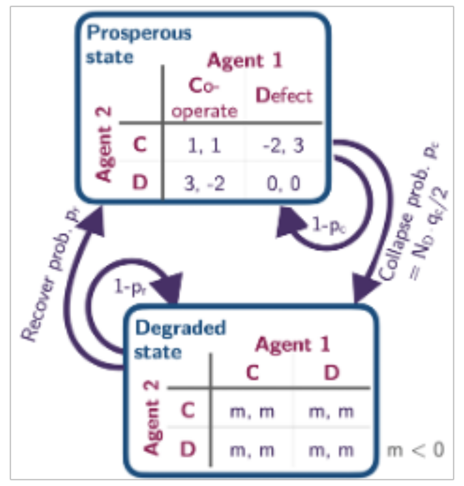
- Bringing the level of understanding from CSS to the richness of MARL by
- Collective Reinforcement Learning Dynamics: mitigating noise in low-dimensional environments

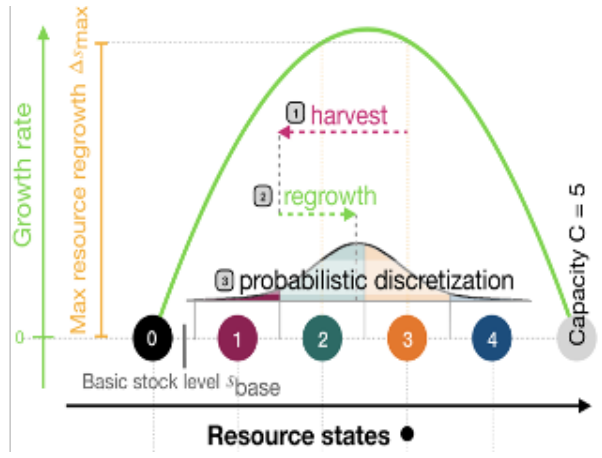
What | Emergent phenomena from cognition in contexts

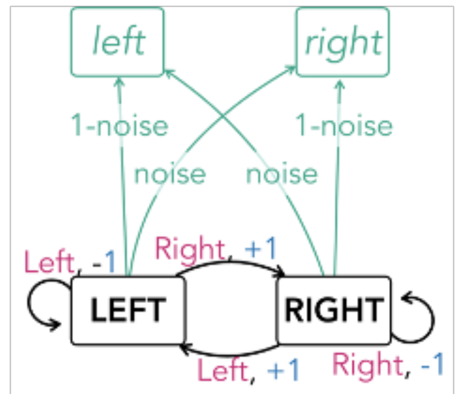



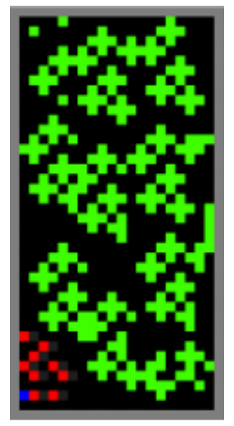
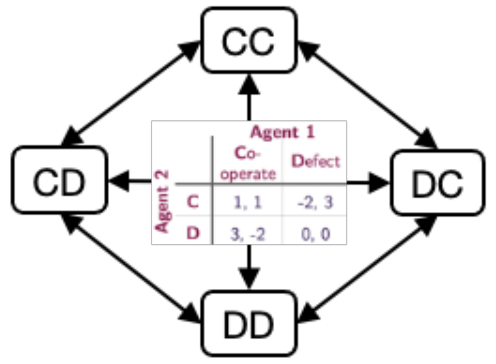








Thank you

wbarfuss@uni-bonn.de
barfusslab.github.io

Cooperative AI

zef
Zentrum für
Entwicklungsforschung
Universität Bonn

TRA
SUSTAINABLE
FUTURES

UNIVERSITÄT BONN