Balancing selfishness and norm conformity can explain human behavior in large-scale Prisoner Dilemma games and pose human groups near criticality

John Realpe-Gómez
Quantum Artificial Intelligence Lab
NASA Ames Research Center
Instituto de Matemáticas Aplicadas
Universidad de Cartagena

Collaborators:
Giulia Andrighetto
Luis Gustavo Nardin
Daniele Villone
Javier Montoya

Realpe-Gómez et al., arXiv:1608.01291
To appear in *Physical Review E*

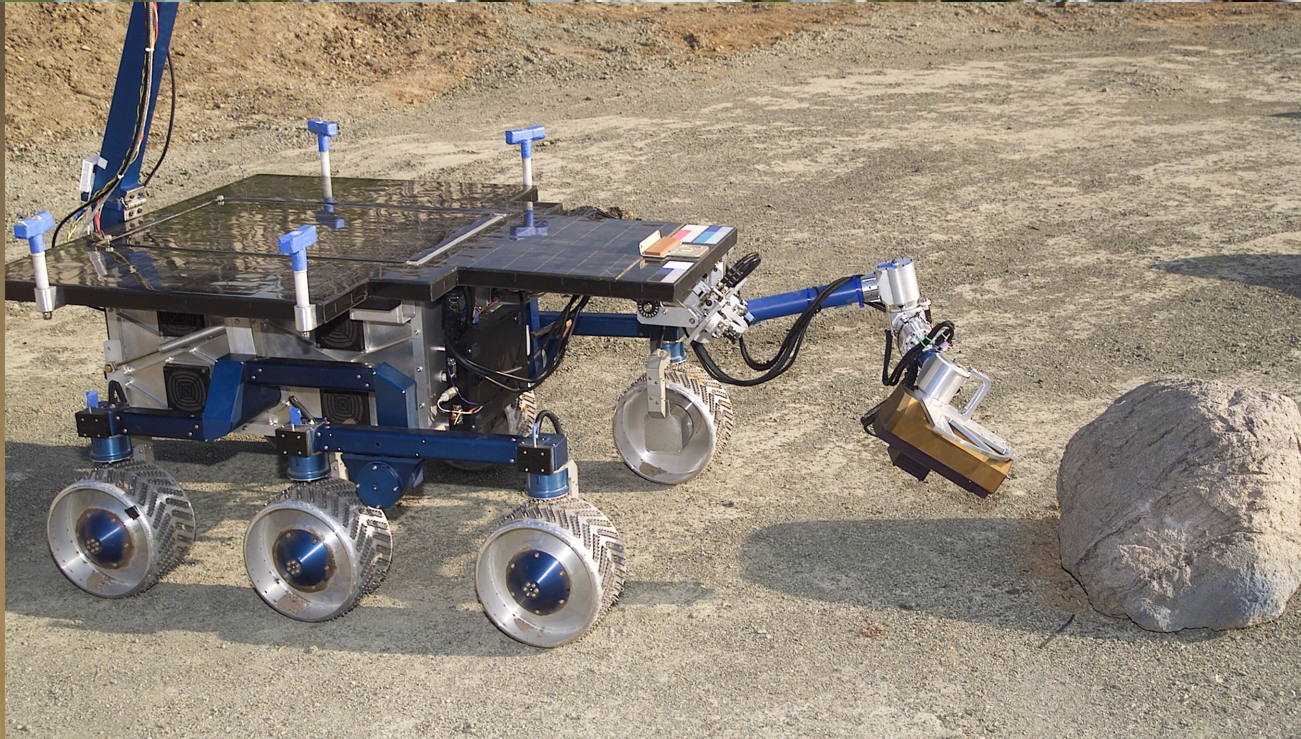
Universidad de Cartagena
Fundada en 1827

Quantil
Bogoá, March 22, 2018


NASA
AMES RESEARCH CENTER
NASA RESEARCH PARK

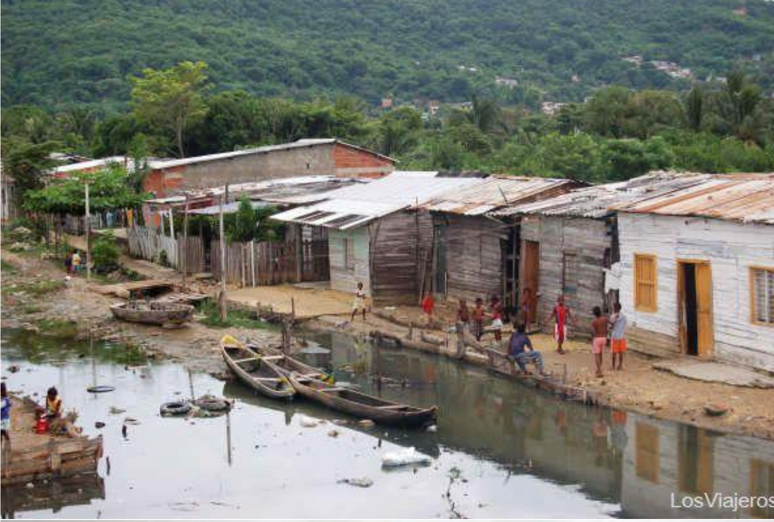

United States


NASA

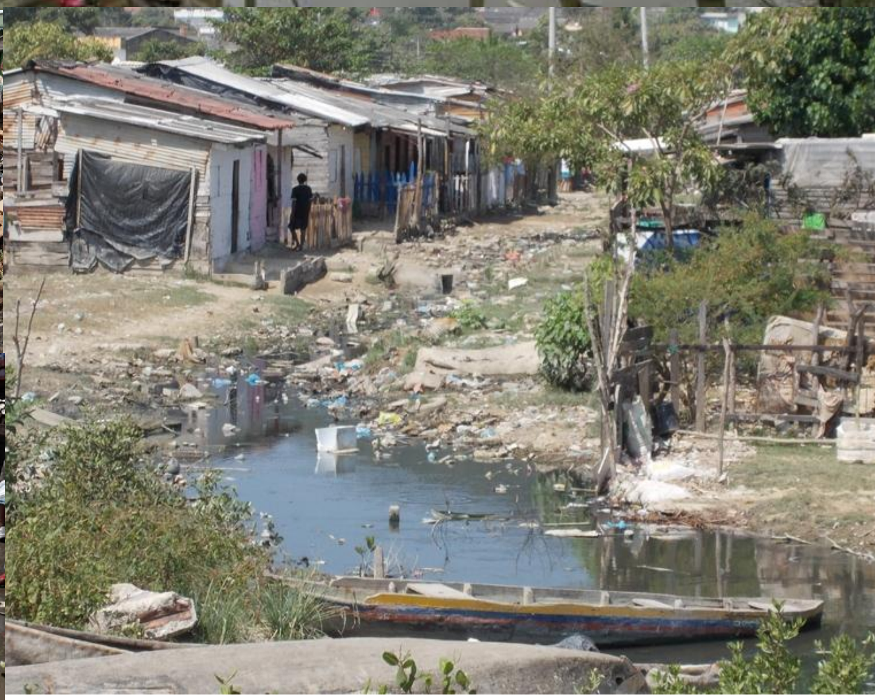







LosViajeros

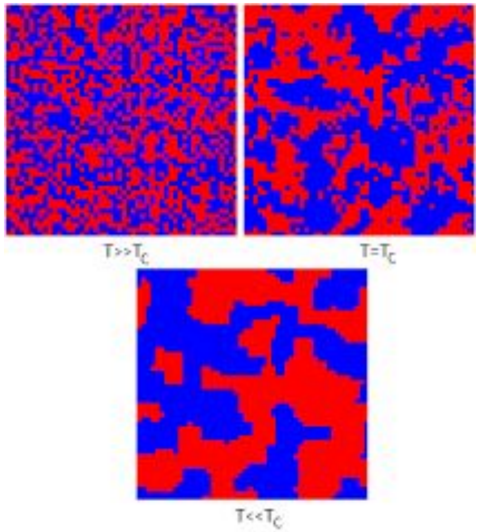




Summary

- Criticality in Complex Systems
- Criticality in Human Social Systems
- Dynamics of social dilemmas and critical behaviour: empirical, simulation and theoretical analysis
- Conclusions

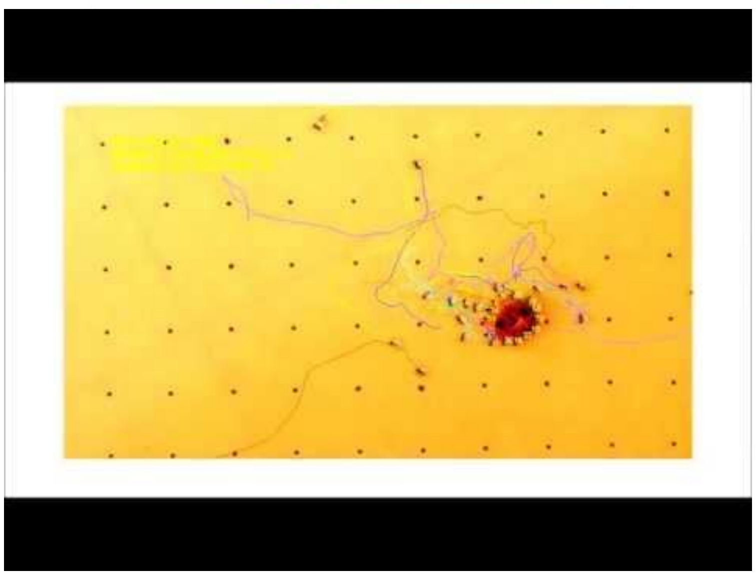
Criticality in Complex Systems



Ising model (applet)

Cavagna *et al.*, *PNAS* (2010)

(animations, top and bottom)



Gelblum *et al.*, *Nat. Comm.* (2015)
(YouTube video)

Effects of criticality

- At a critical point, a system has long-range correlations (classical thermodynamics).
- Close to a critical point, a system is able to explore more possible configurations.
- A biological system near criticality maximizes the fitness and shows resilience (Hidalgo *et al.*, *PNAS* 2014).
- Could it be valid also for Human Social Systems?

Why criticality?

How does a collective biological system reach a critical configuration?

Several mechanisms have been proposed:

- Criticality stems from the optimal balance between individuality and conformism (Gelblum *et al.*, *Nat. Comm.* 2015);
- Criticality origins from the mutual adaptation of agents inferring their peers' behaviour (Hidalgo *et al.*, *PNAS* 2014).

Criticality in Human Social Systems

First experimental evidence of criticality when humans play Prisoners' Dilemma: Realpe-Gómez, *et al.* (2017).

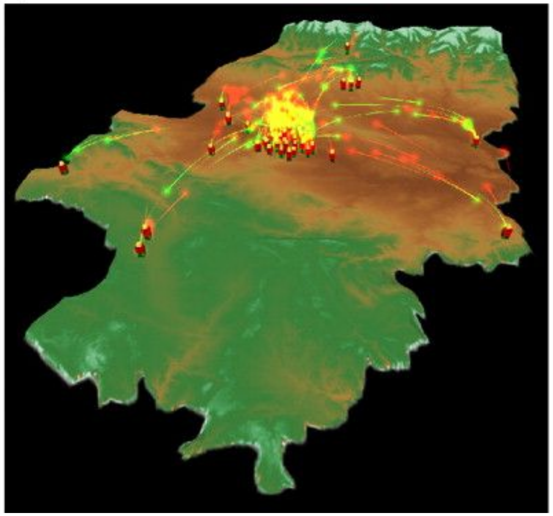
Mechanisms proposed:

- balancing individual and norm-based considerations (cf. Gelblum, *et al.*, 2015);
- learning from peers' behaviour (cf. Hidalgo *et al.*, 2014).

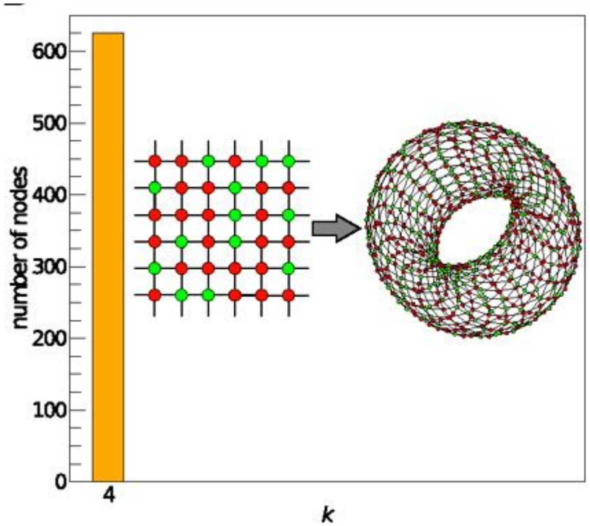
Experimental setup analysed: Large-scale Prisoner's Dilemma Game in Gracia-Lázaro *et al.*, *PNAS* (2012)

Human and Social Dilemmas

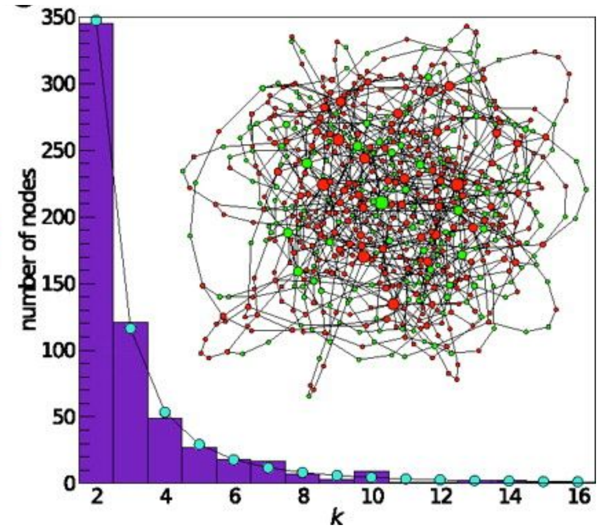
Aragón



Lattice



Heterogeneous



Main experimental observations

1. Lattices or networks do *not* support cooperation.
2. People display *Moody Conditional Cooperation* (MCC), i.e., when deciding to cooperate individuals are responsive to the behavior of others, but only if they have cooperated themselves.
3. People do *not* take into account the earnings of their neighbors.
4. Cooperation *can* be sustained in dynamic networks.

García-Lázaro et al PNAS 2012; Sanchez JSTAT 2018

Humans and Social Dilemmas

In experiments (again PDG) conducted by Grujić *et al.*, *PloS One* (2010), three kinds of players have been identified:

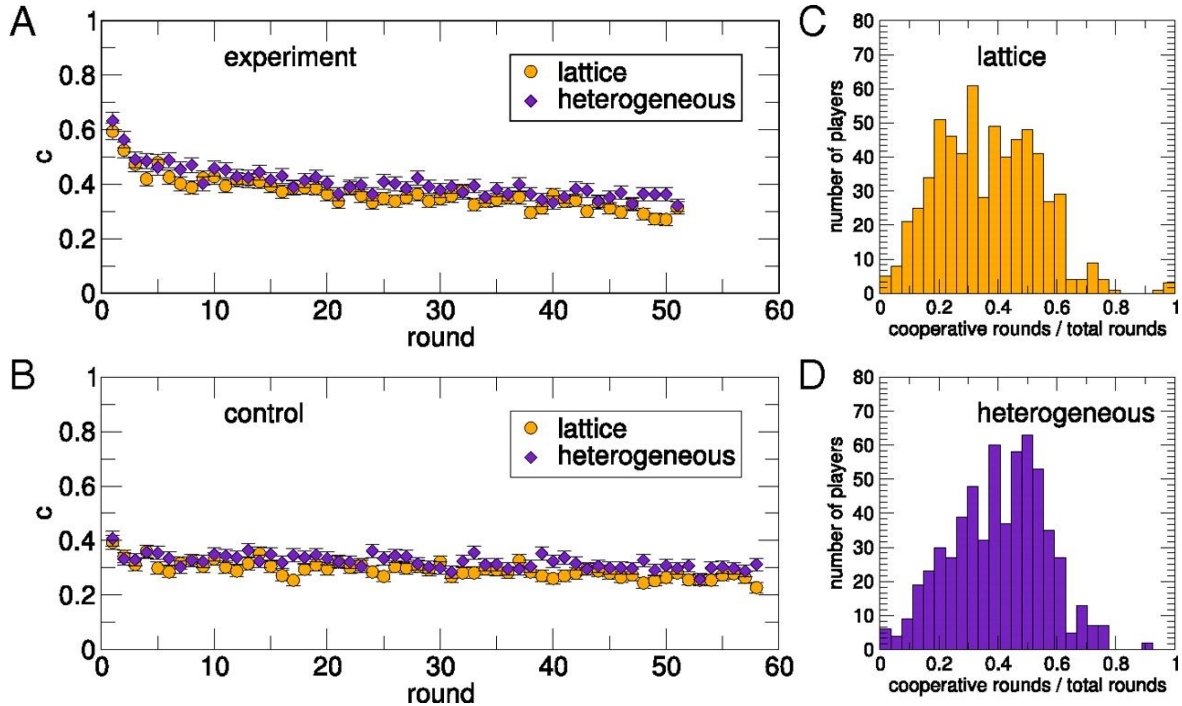
- absolute cooperators (~5%),
- absolute defectors (~30%),
- agents which respond to the cooperation they observe in a reciprocal manner, the so-called Moody Conditional Cooperators (MCC, ~65%).

The MCCs are the only players able to adapt their behaviour to the actions of others and the social norms ruling the environment.

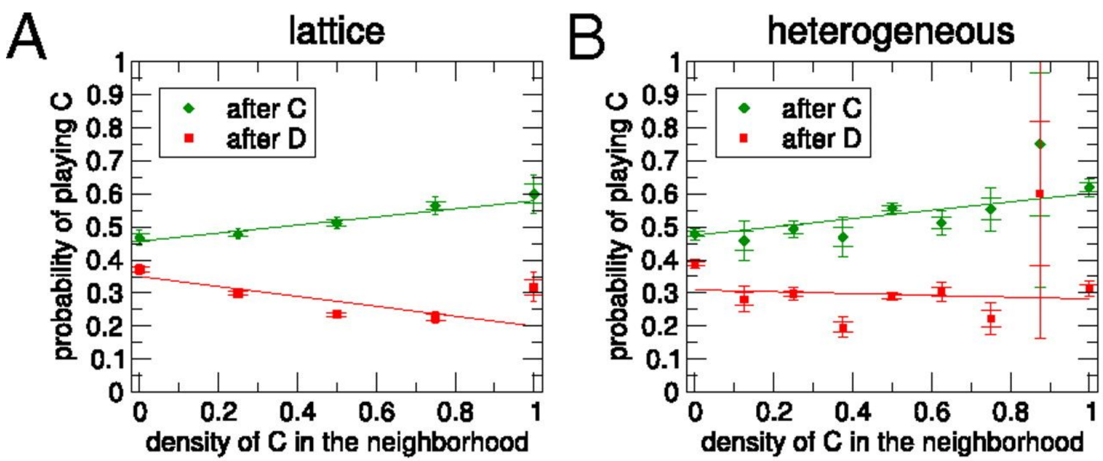
The analysis of criticality in Realpe-Gómez *et al.* has been based on a representative agent similar to MCCs

Some experimental results with 625 human subjects

Global feature:
average cooperation



Local feature:
moody conditional
cooperation



García-Lázaro et al PNAS 2012

Towards a more realistic modeling of human behavior

Agents weight "utility" of selfish and prosocial behavior

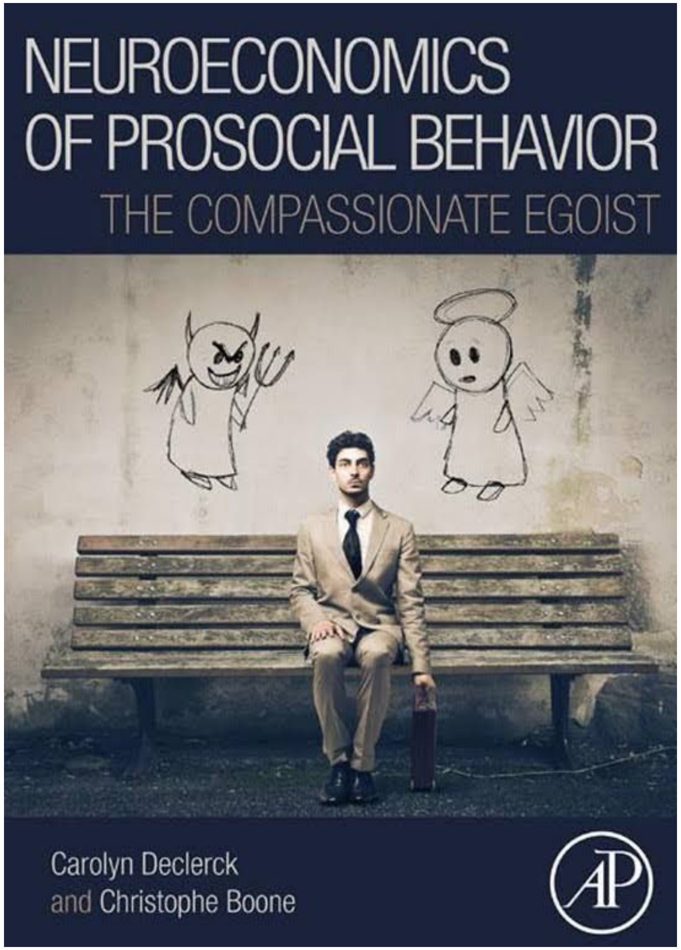
$$\Delta U_i(t) = \Delta I_i(t) + h \Delta N_i(t)$$

Agents have decaying memory of performance.
Drive to cooperate given by:

$$D_i(t+1) = (1-\alpha)\, D_i(t) + \Delta U_i(t)$$

Bounded rationality: agents make "mistakes".
Probability to cooperate given by

$$x_i(t+1) = \frac{1}{1 + e^{-\beta D_i(t+1)}}$$



Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Towards a more realistic modeling of human behavior

Payoffs

	C	D
C	R	S
D	T	P

Experimental values for (weak) Prisoner's Dilemma

		ELECCIÓN DE UN VECINO	
SU ELECCIÓN			
		7	0
		10	0

Individual drive ($\Delta I_C = R - T$ and $\Delta I_D = S - P$)

$$\Delta I_i(t) = (\Delta I_C - \Delta I_D)\frac{1}{K}\sum_{j\in\partial i} s_j(t) + \Delta I_D$$

Normative drive

$$\Delta N_i(t) = w_C[2s_i(t) - 1] + w_O\frac{1}{K}\sum_{j\in\partial i} s_j(t) + w_I s_i(t)\frac{1}{K}\sum_{j\in\partial i} s_j(t)$$

Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Towards a more realistic modeling of human behavior

	Assumption	Description	Representation
1st block	Bounded rationality	Agents do not always play the optimal strategy	β in Eq. (1)
	Belief learning	Agents learn from what could have *potentially* happened	Eq. (2)
	Reinforcement learning	Agents learn from what *actually* happened	Eq. (2)
	Memory decay	Agents give more relevance to recent events	α in Eq. (2)
	Selfishness	Agents base their decisions on self-regarding considerations	ΔI_C, ΔI_D, Eqs. (3) and (4)
2nd block	Norm conformity:	Agents base their decisions *also* on social norms	h in Eqs. (3) and (5)
	- Self-consistency	Agents are consistent with own beliefs and self-ascribed norms	w_C in Eq. (5)
	- Social influence	Norm compliance increases with the number of compliant peers	w_O in Eq. (5)
	- Moody conditional coop.	Social influence is stronger if aligned with self-consistency	w_I in Eq. (5)
3rd block	Slow adaptation	Adaptation happens over several individual strategic choices	Eqs. (9) and (10)
	No network reciprocity	Interaction structure does not significantly influence behavior	Eqs. (11) and (12)

Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Further empirically motivated simplifying assumptions

Absence of network reciprocity (mean field approximation)

$$\sum_{j \in \partial i} x_j \approx xK$$

Slow adaptation (adiabatic approximation)

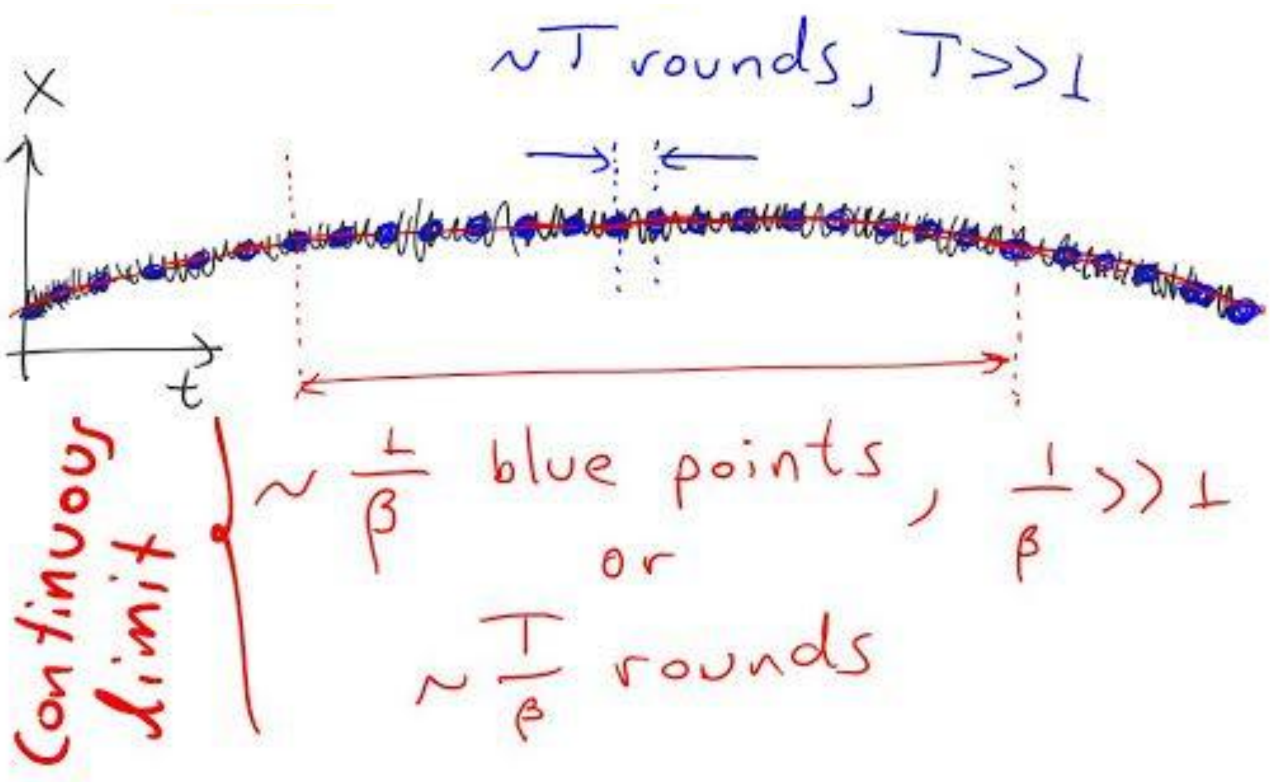
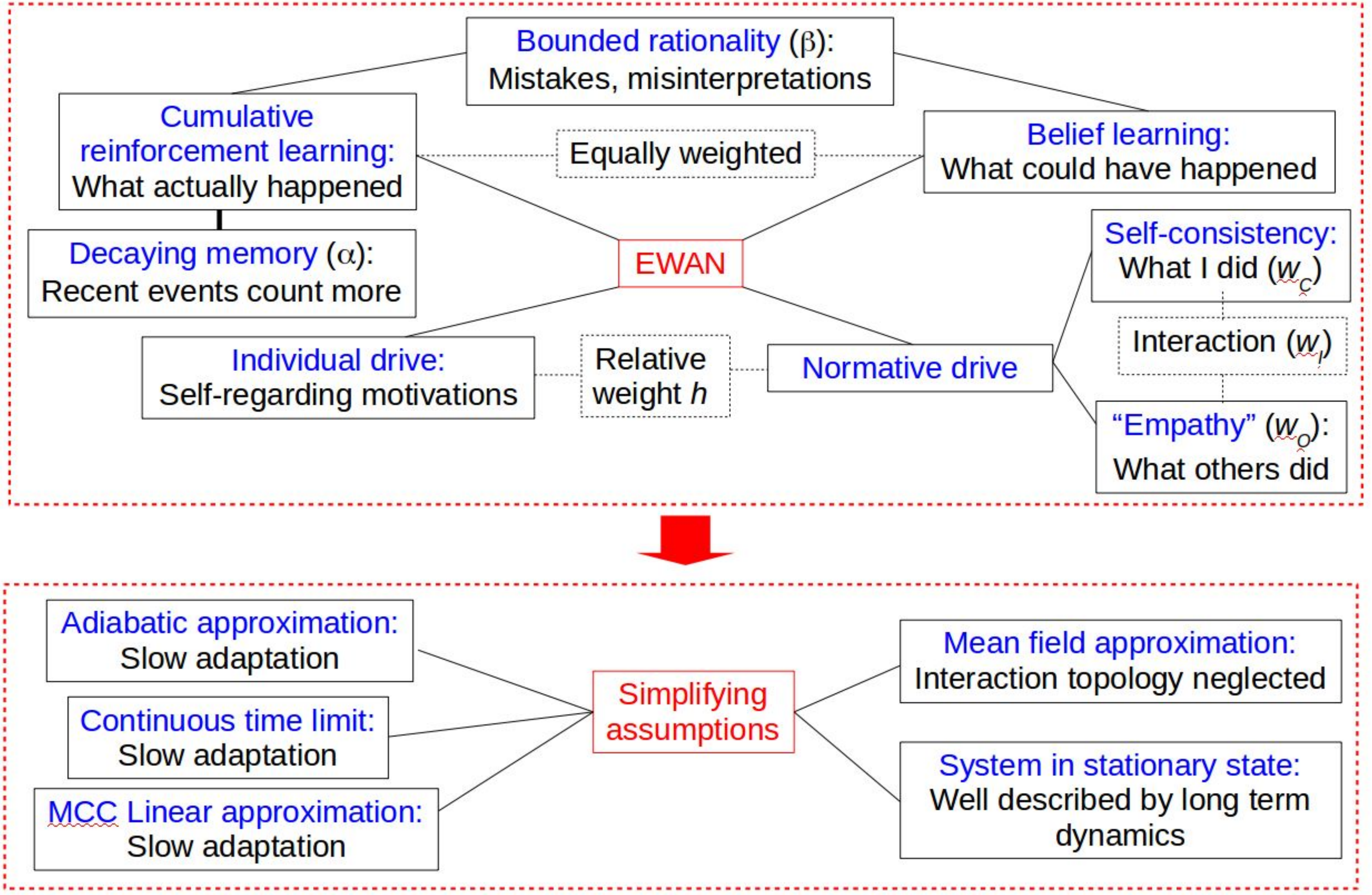


Diagram of model assumptions



Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Single-representative agent model and long-term dynamics

Final deterministic (adiabatic approx.) single representative agent (mean field approx.) dynamical equation given by:

$$x(t+1) = \frac{x(t)^{1-\alpha}}{x(t)^{1-\alpha} + [1 - x(t)]^{1-\alpha}\, e^{-\beta \overline{\Delta U}[x(t)]}}$$

where effective utility function in terms of effective parameters is given by:

$$\overline{\Delta U}[x] = aK\, x^2 + (bK + 2h)\, x - h$$

Long-term dynamics can be characterized by fixed points of the equation

$$x(t+1) = x(t) = x$$

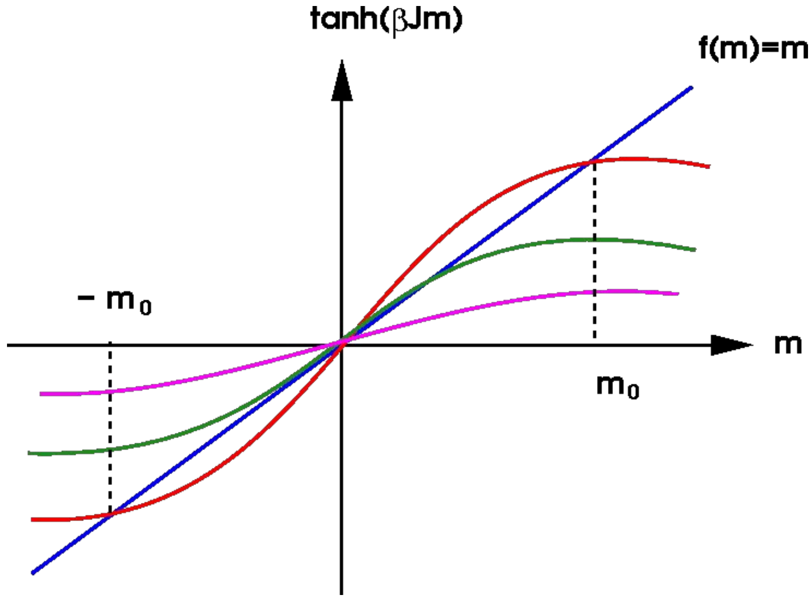
which yields

$$x = f(x) \quad \text{with} \quad f(x) = \frac{1}{2} + \frac{1}{2}\tanh\left[A(x - x_0)^2 + y_0\right]$$

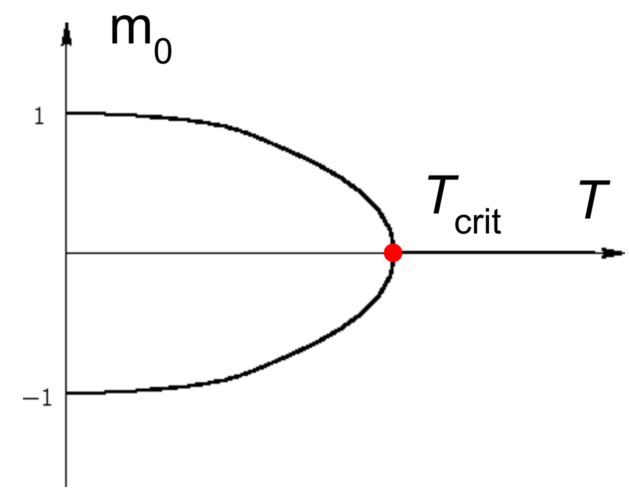
Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Fixed points, phase transitions, and criticality

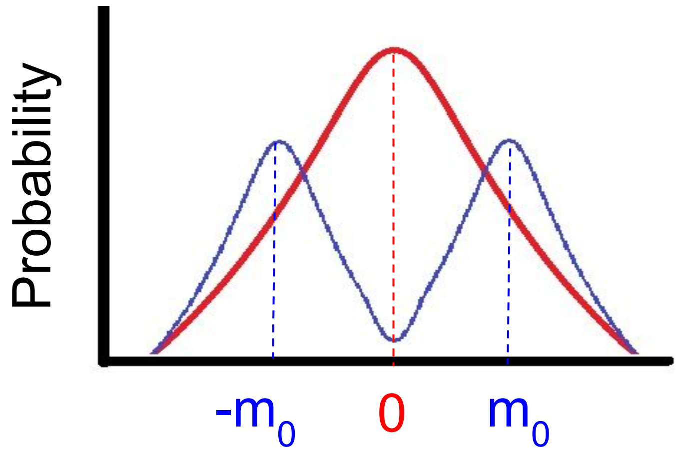
Fixed-point equation for magnetic systems



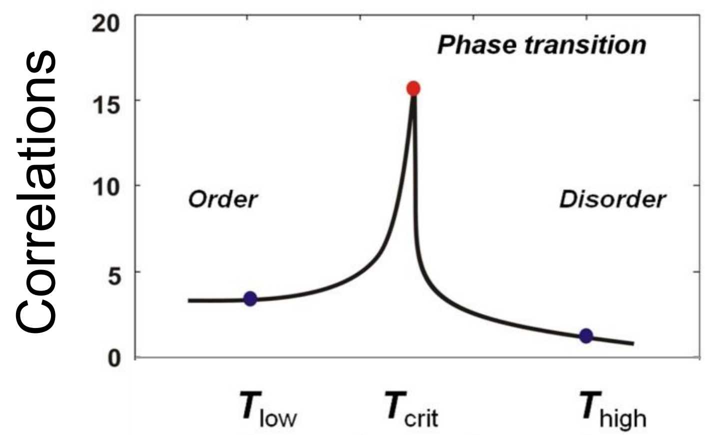
Phase diagram ($T = 1 / \beta$)



Phase transition from mono- to bi-modal



Susceptibility, responsiveness



Fixed points, phase transitions, and criticality

EWAN model's fixed points

$$x = f(x) \quad \text{with}$$

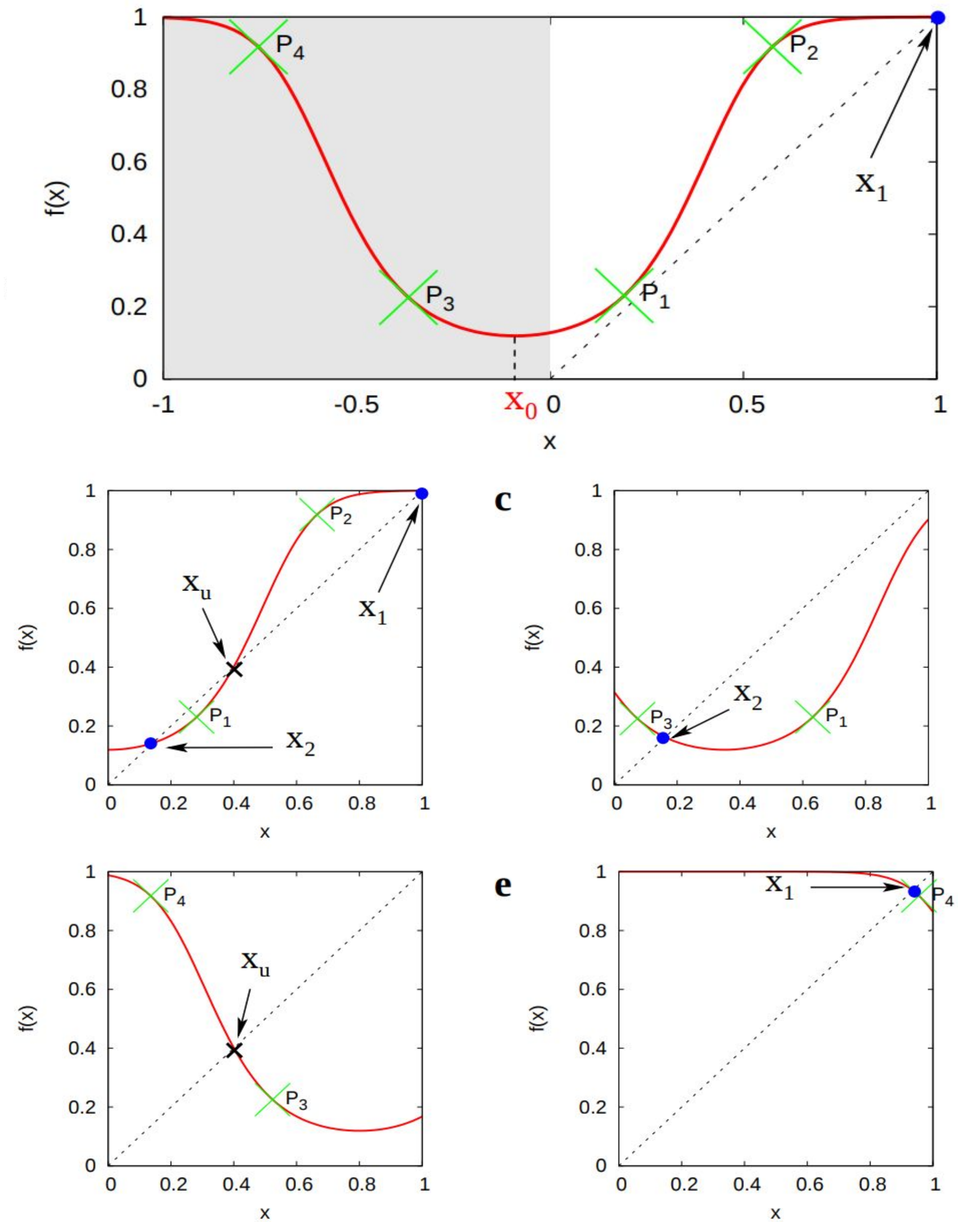
$$f(x) = \frac{1}{2} + \frac{1}{2}\tanh\left[A(x - x_0)^2 + y_0\right]$$

If $w_I = 0$, magnetic-like system

$$m = \tanh\left[\beta(J_{\text{eff}} m + H_{\text{eff}})\right]$$

with

$$J_{\text{eff}} = (hw_O + \Delta I_C + 2h)/4\alpha$$
$$H_{\text{eff}} = (hw_O + \Delta I_C)/4\alpha$$



Susceptibility, responsiveness

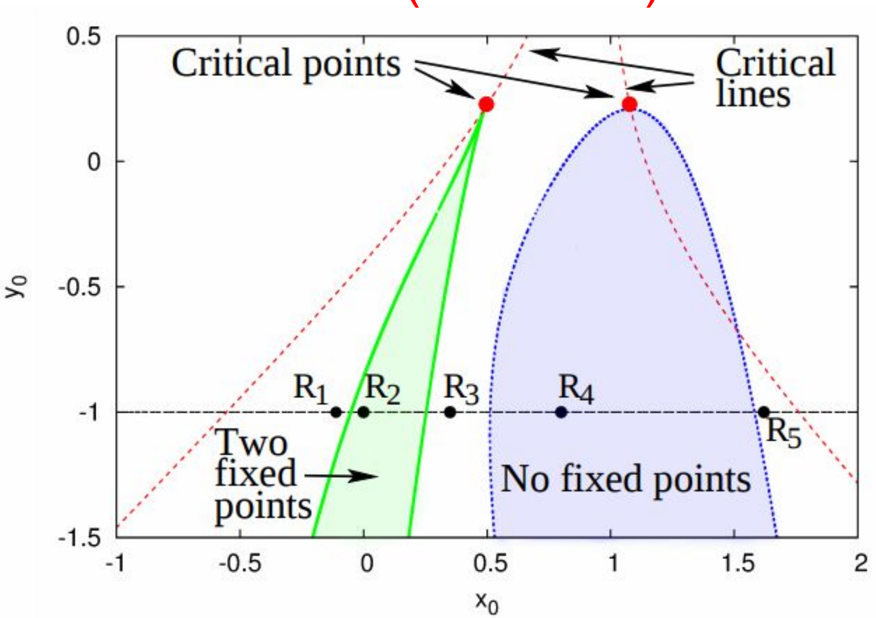
$$\chi = \frac{\partial x}{\partial \theta} \propto \frac{1}{\theta} = \frac{1}{A_c^* - A} \xrightarrow[A \to A_c^*]{} \infty,$$

with $\theta = A_c^* - A$

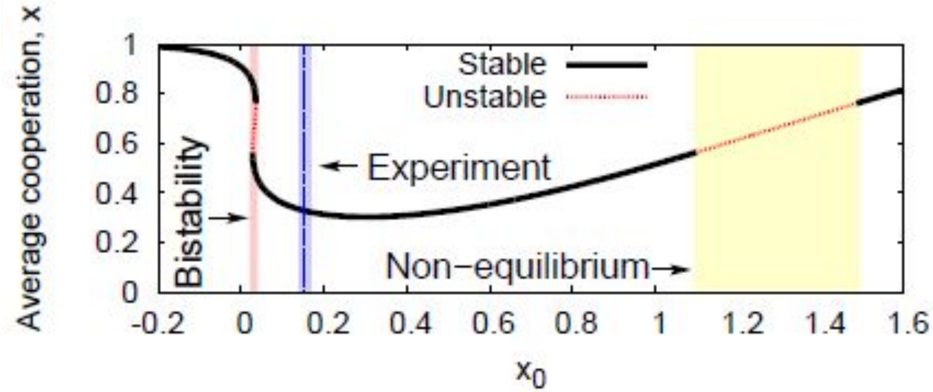
Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Phase diagram and location of experimental human groups

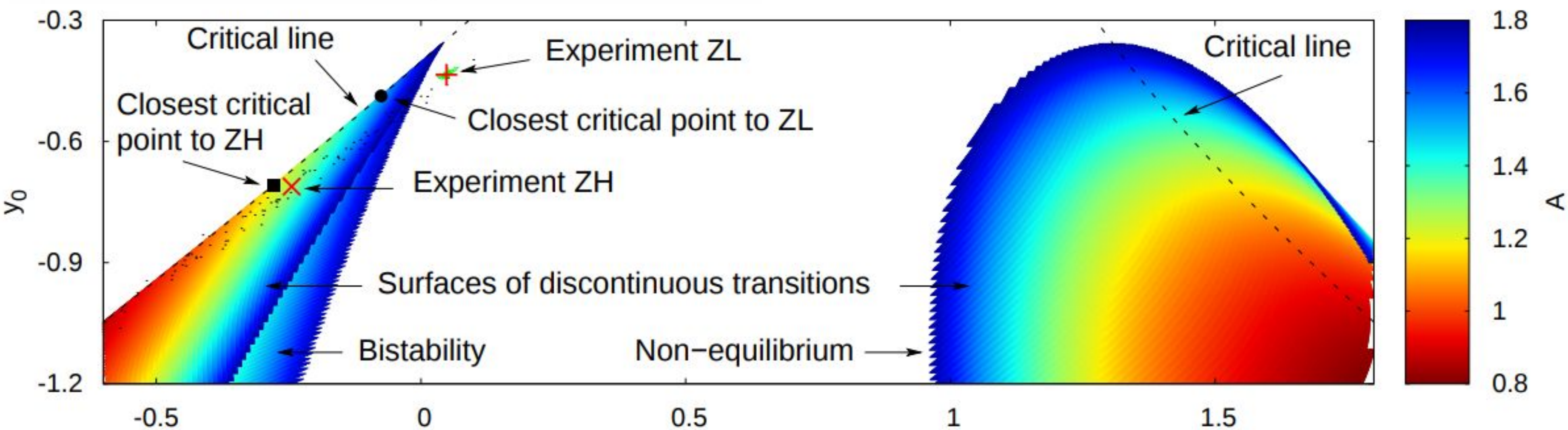
Parameter *A* (curvature) constant



Global cooperation Vs. x_0 (minimum)



Color map for different values of *A*



Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Moody conditional cooperation and EWAN model

Dynamical equation can be interpreted as:

$$x(t+1) = P\left(C, t+1 | s, n, x, t\right)$$

If there is only one fixed point x_1, there is no dependency on the history, i.e. on x, and at the stationary state (fixed point) we have

$$P\left(C|s,n\right) = \frac{1}{1 + y_1^{1-\alpha} e^{-\beta \Delta U(s,n)}}$$

where $y_1 = \left(1 - x_1\right)/x_1$

When rationality parameter β is small, we can do a linear expansion:

$$P\left(C|s,n\right) = m_s n/K + r_s$$

Where slopes and intercepts are given by

$$m_s = \beta K J(\alpha)(as + b),$$
$$r_s = I(\alpha) + \beta J(\alpha)\left[h(2s-1)\right]$$

$$I(\alpha) \equiv \frac{1}{1 + y_1^{1-\alpha}},$$
$$J(\alpha) \equiv \frac{y_1^{1-\alpha}}{\left(1 + y_1^{1-\alpha}\right)^2}.$$

Realpe-Gómez et al. arXiv:1608.01291, to appear in *Physical Review E*

Bayesian parameter inference from experimental data

Joint distribution of "true" deterministic trajectory and noisy observed one:

$$\mathcal{P}[\mathbf{x}(0:T), \mathbf{x}_{\text{obs}}(1:T)|\Theta] = \mathcal{P}_0[x(0)] \prod_{t=1}^{T} \mathcal{P}_{\text{obs}}[x_{\text{obs}}(t)|x(t)]\, \mathcal{P}_{\text{dyn}}[x(t)|x(t-1)|\Theta],$$

where

$$\mathcal{P}_{\text{dyn}}[x(t)|x(t-1)|\Theta] = \delta[x(t) - x(t-1)] \quad \text{(Dirac delta function)}$$

$$\mathcal{P}_{\text{obs}}[x_{\text{obs}}(t)|x(t)] = \mathcal{N}[x_{\text{obs}}(t); x(t), \sigma]$$

Parameter inference: Compute posterior

$$\mathcal{P}_\theta[\Theta|\mathbf{x}_{\text{obs}}(1:T)] \propto \mathcal{P}[\mathbf{x}_{\text{obs}}(1:T)|\Theta]\mathcal{P}_{\text{prior}}[\Theta],$$

where $\Theta = \mathcal{O} \equiv (m_C, m_D, r_C, r_D).$

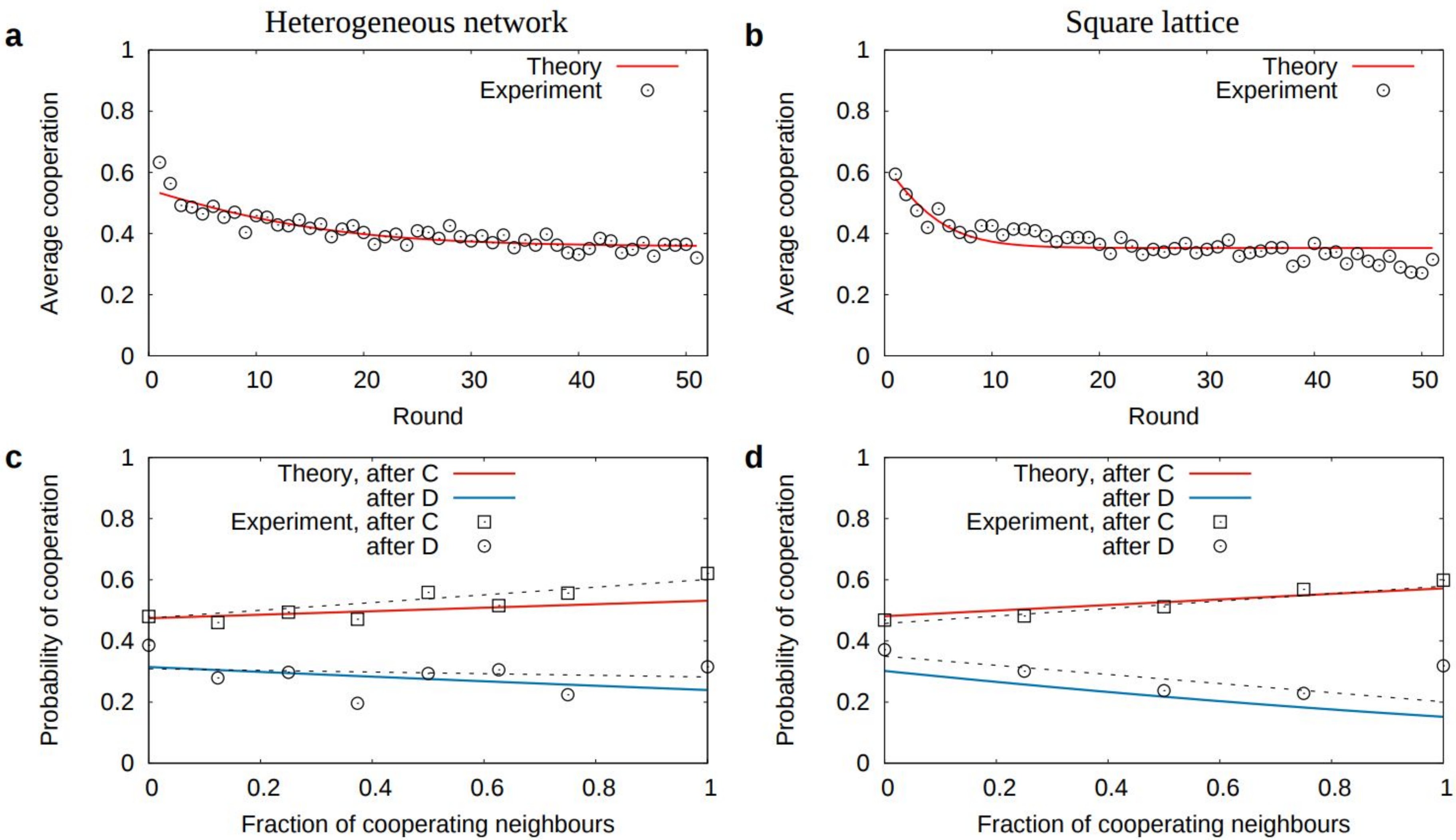
Prior was chosen from values allowed by experimental error, i.e.

$$\mathcal{P}_{\text{prior}}[\Theta], = \text{Uniform in } [O^* - \zeta\delta O^*, O^* + \zeta\delta O^*].$$

$\zeta = 1.28$ yields 90% credible interval. $\zeta = 1.96$ yields 97.5% credible interval.

Reproducing experimental results with EWAN model

Experiments with 625 humans



Impact of EWAN model parameters

We can also describe the MCC linear trend in terms of mean intercept r and gap G between intercepts

$$r = \frac{1}{2}(r_C + r_D) = I(\alpha),$$
$$G = r_C - r_D = 2\beta h w_C J(\alpha),$$

as well as the difference and ratio between slopes

$$m_C - m_D = \beta\, a K J(\alpha),$$
$$\frac{m_C}{m_D} = \frac{\beta\, a + \beta\, b}{\beta\, b}.$$

So,

- If "mood parameter" w_C = 0, then gap vanishes, G = 0. **Not observed.**
- If "MCC parameter" w_I = 0, so a = 0, then slopes equal, $m_C = m_D$. **Not observed.** Moreover, w_I generates non-equilibrium phenomena.
- If "peer pressure" parameter w_O = 0, slope m_D always negative. **Observed empirically, yet w_O was required for good fit.**

Final remarks

- The network structure has much less influence than the mere number of neighbours: this is a typical feature of critical phenomena (universality classes).
- Social norm driven behavior (as MCCs' behaviour) poises the system to a critical point.
- Further studies are still needed (of course!): in particular, new laboratory experiments designed to test directly for criticality, as well as the analysis of finite size effects, are necessary to reach more solid conclusions.
- References: Vilone, Andrighetto, Realpe-Gómez, *Studies in Computational Intelligence*, **689** (2018);
Vilone, Andrighetto, Realpe-Gómez, *in preparation*.

THANK YOU!