石头－剪刀－布的统计物理学
STAT-PHYS OF ROCK-PAPER-SCISSORS

周海军
中国科学院理论物理研究所

合作者:
王志坚（浙江大学实验社会学实验室）、**许彬**（浙江工商大学公共管理学院）

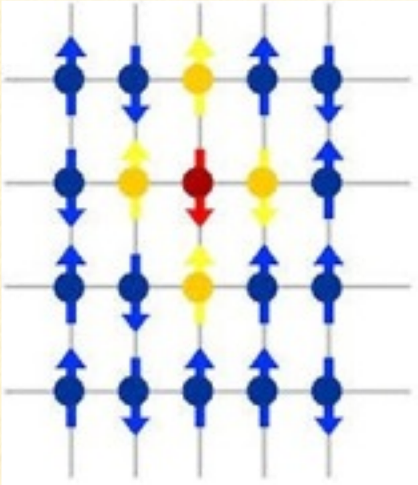
Zhijian Wang, Bin Xu, Hai-Jun Zhou*,
Social cycling and conditional responses in the Rock-Paper-Scissors game,
arXiv:1404.5199 (April 21, 2014)

STAT-PHYSICS AS A BRIDGE

macroscopic phenomena

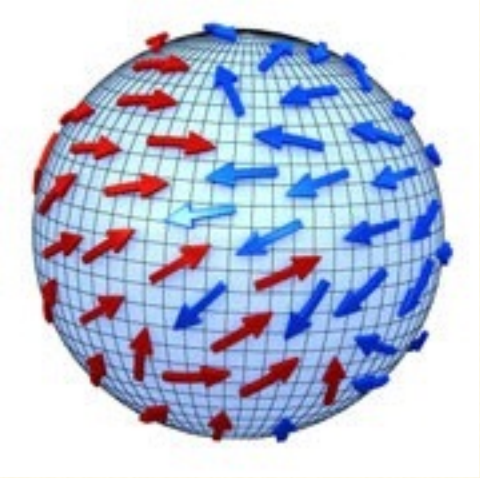
* infer micro-interactions from macro-properties (the **inverse** problem)

* predict macro-properties from micro-interactions (the **direct** problem)



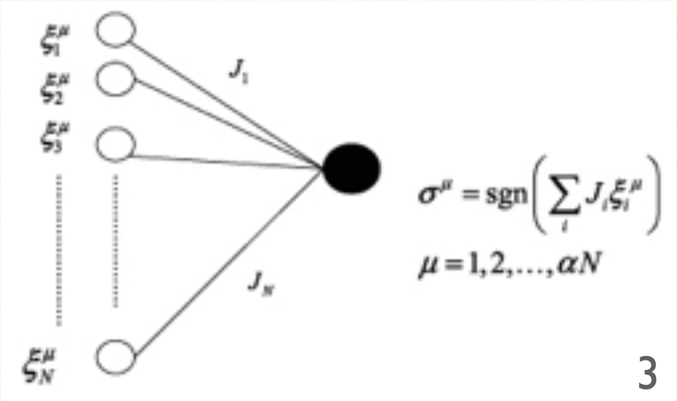
microscopic interactions

我们研究小组



- **自旋玻璃**：平均场理论、消息传播算法、组合优化、约束满足、复杂网络 (小册子即将出版）

- **博弈动力学**：决策、非平衡演化、优化
- **学习动力学**：神经网络、学习微观机制

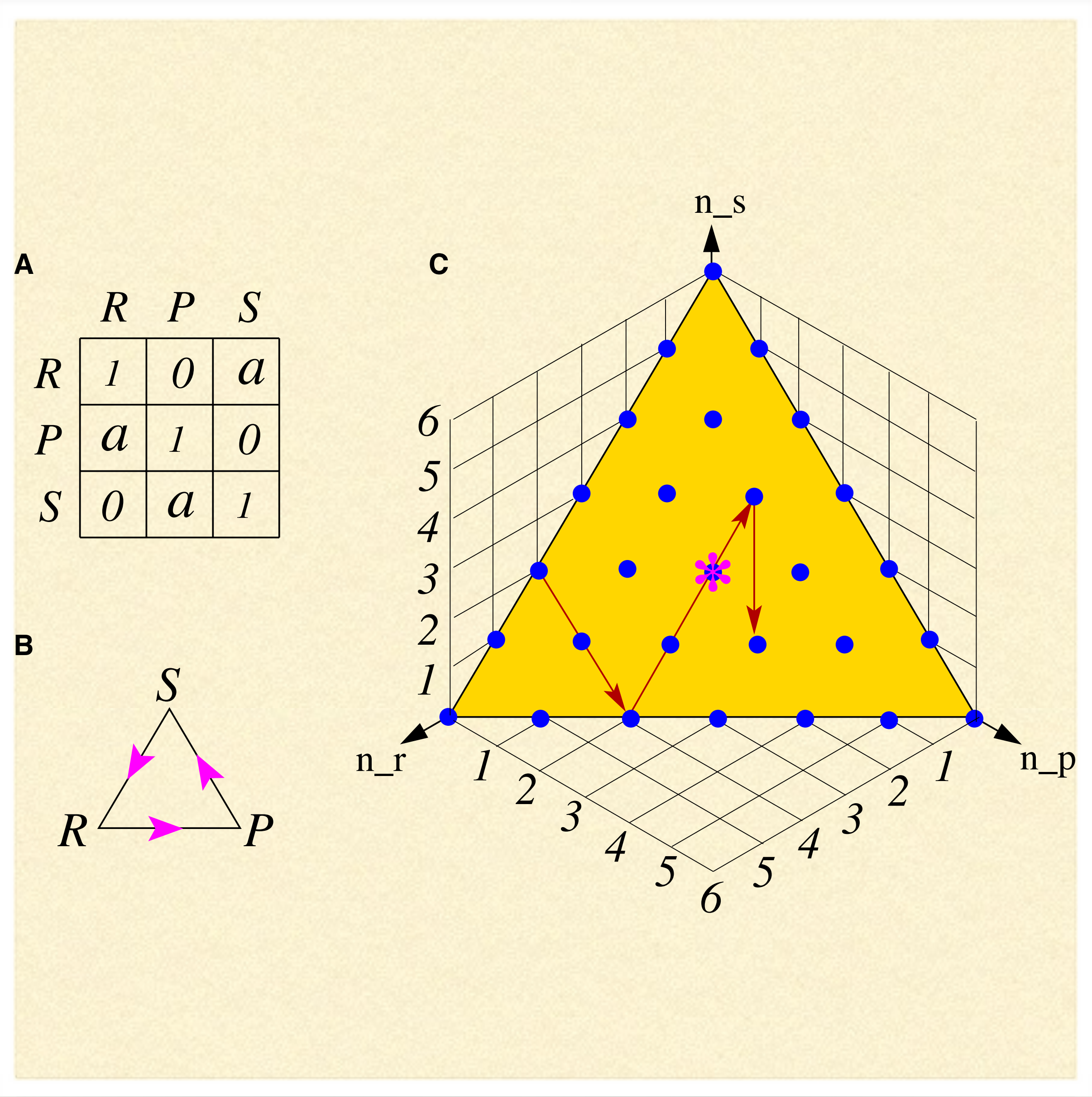


面向交叉学科，面向复杂系统

OUTLINE

- background
- experimental setup
- experimental observations
- theoretical modeling
- discussions

ROCK-PAPER-SCISSORS GAME



- Basic model of non-cooperative strategic interactions
- Only single parameter: payoff ***a*** of winning action
- For a population of ***N*** players, social state denoted as (n_r, n_p, n_s)
- social state evolution

EXPERIMENTAL SET-UP

- Finite population, N=6
- Random pairwise-matching: at each game round, a player competes with a random opponent
- win: *a* points;
  tie: *1* point
  lose: *0* point
- Game repeats 300 rounds
- Feedback information: own payoff; own action; opponent's action; own accumulate payoff

payoff parameter *a:*

$a = 1.1$ (11 populations, NE non-stable)
$a = 2$ (12 populations, NE neutral)
$a = 4$ (12 populations, NE stable)
$a = 9$ (12 populations, NE stable)
$a = 100$ (12 populations, NE stable)

CLASSICAL GAME THEORY
VS
EVOLUTIONARY GAME THEORY

CGT: Nash (1950)

- complete rationality.
- mixed-strategy Nash equilibrium.

$$\begin{array}{l} \frac{1}{3} - \mathrm{R} \\ \frac{1}{3} - \mathrm{P} \\ \frac{1}{3} - \mathrm{S} \end{array}$$

for each player
at each game round

- social dynamics trivial (detailed balance).

EGT: Maynard Smith (1973)

- bounded rationality
- microscopic evolutionary mechanisms or learning mechanisms
- breaking of detailed balance.
- individual- and/or social-level cycling.

OUR QUESTION

Resolution of CGT—EGT debate in human subjects systems rather challenging:

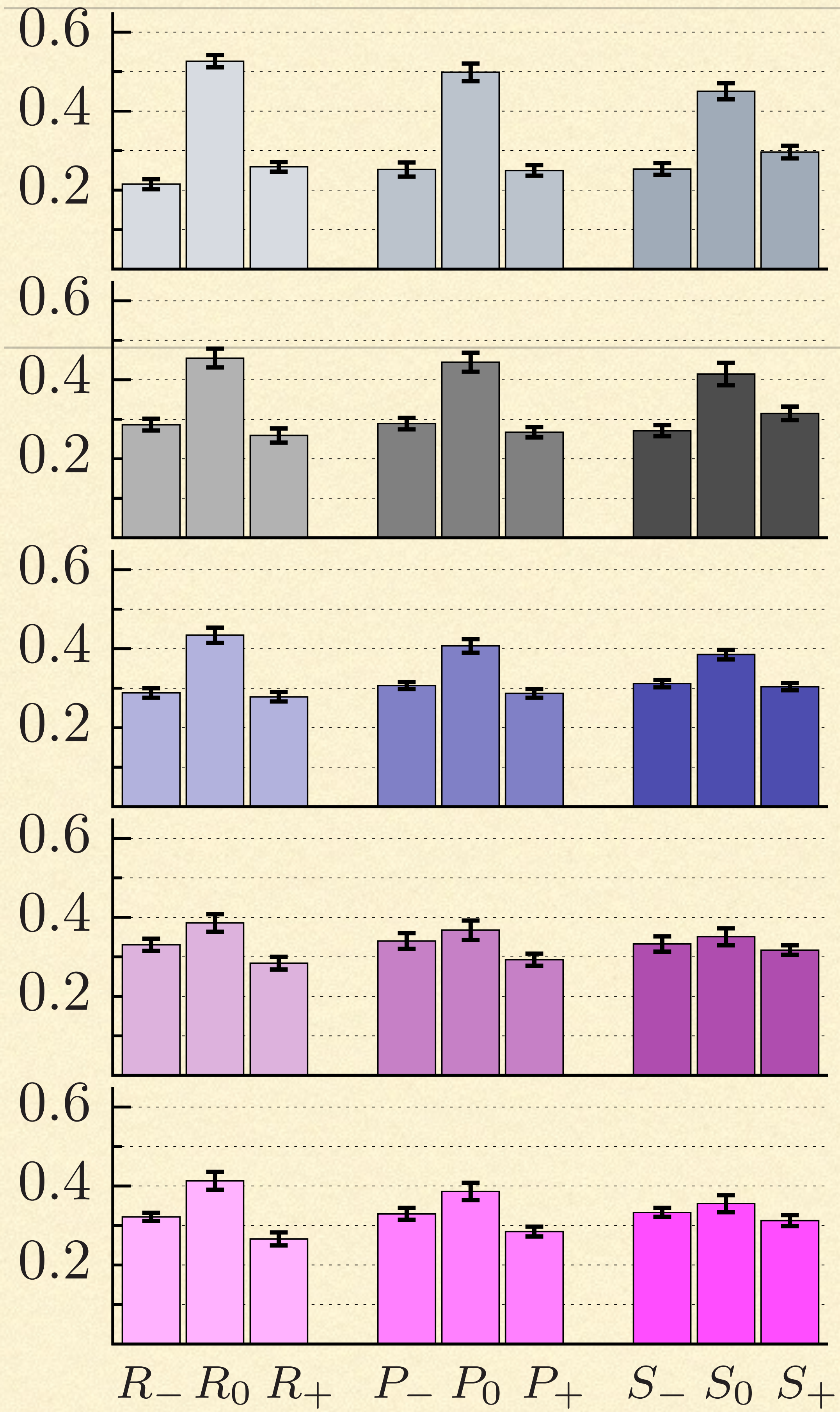
1) Experimental data very noisy;

2) Experiments can't take sufficiently long;

3) Game theorists are mainly mathematicians (they love paper work) …

How humans make decisions in non-cooperative situations under only partial information?

The finite-population RPS game.



BEHAVIOR OF INDIVIDUAL PLAYER

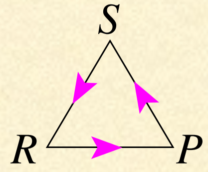
- Individual players change actions frequently:

  R: 0.36 ± 0.08 (mean $\pm$ s.d.)
  P: 0.33 ± 0.07
  S: 0.32 ± 0.06

  consistent with Nash equilibrium

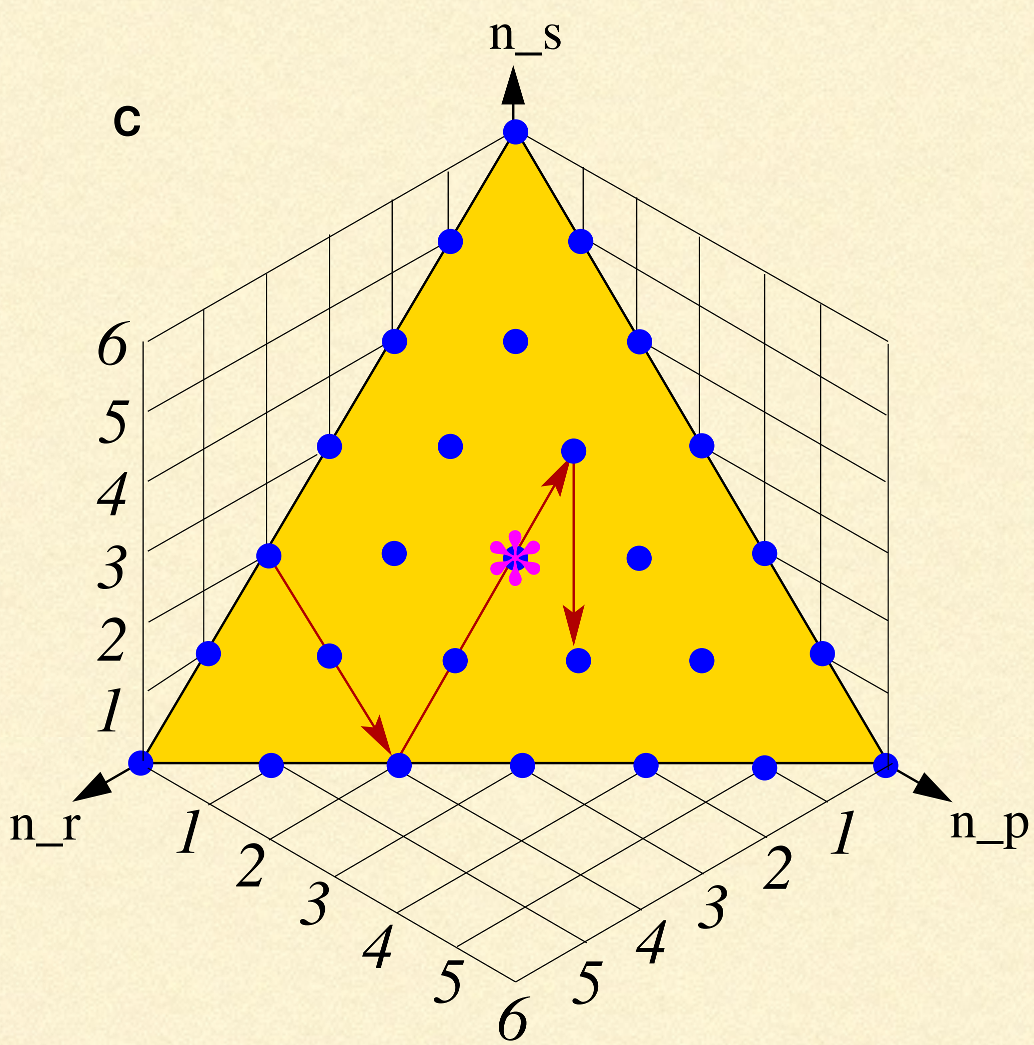
- Inertial effect: more likely to choose same action than to shift action either clock wisely (-) or counter-clock wisely (+)

  different with Nash equilibrium



- No individual-level cycling.

COLLECTIVE BEHAVIOR: ROTATION AROUND CENTROID

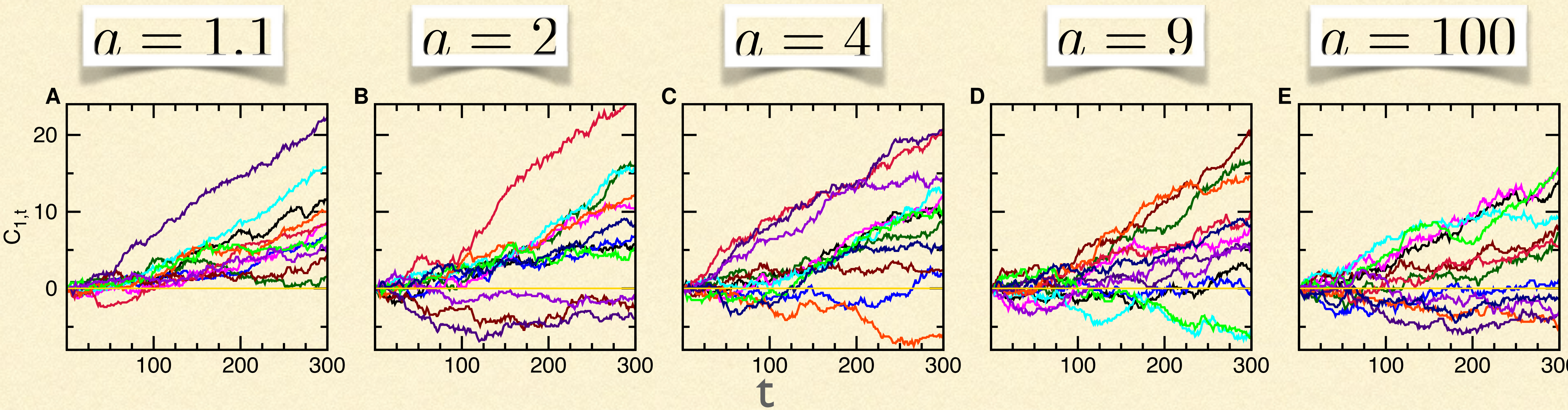


- rotation number:

$$C_{t_0,t_1} \equiv \sum_{t=t_0}^{t_1-1} \frac{\theta(t)}{2\pi}$$

- rotation frequency:

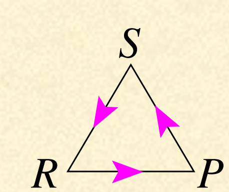
$$f_{t_0,t_1} \equiv \frac{C_{t_0,t_1}}{t_1 - t_0}$$



$f_{1,300}$

mean	0.031	0.027	0.031	0.022	0.018
s.d.	0.019	0.029	0.026	0.027	0.025
s.e.m.	0.006	0.008	0.008	0.008	0.007

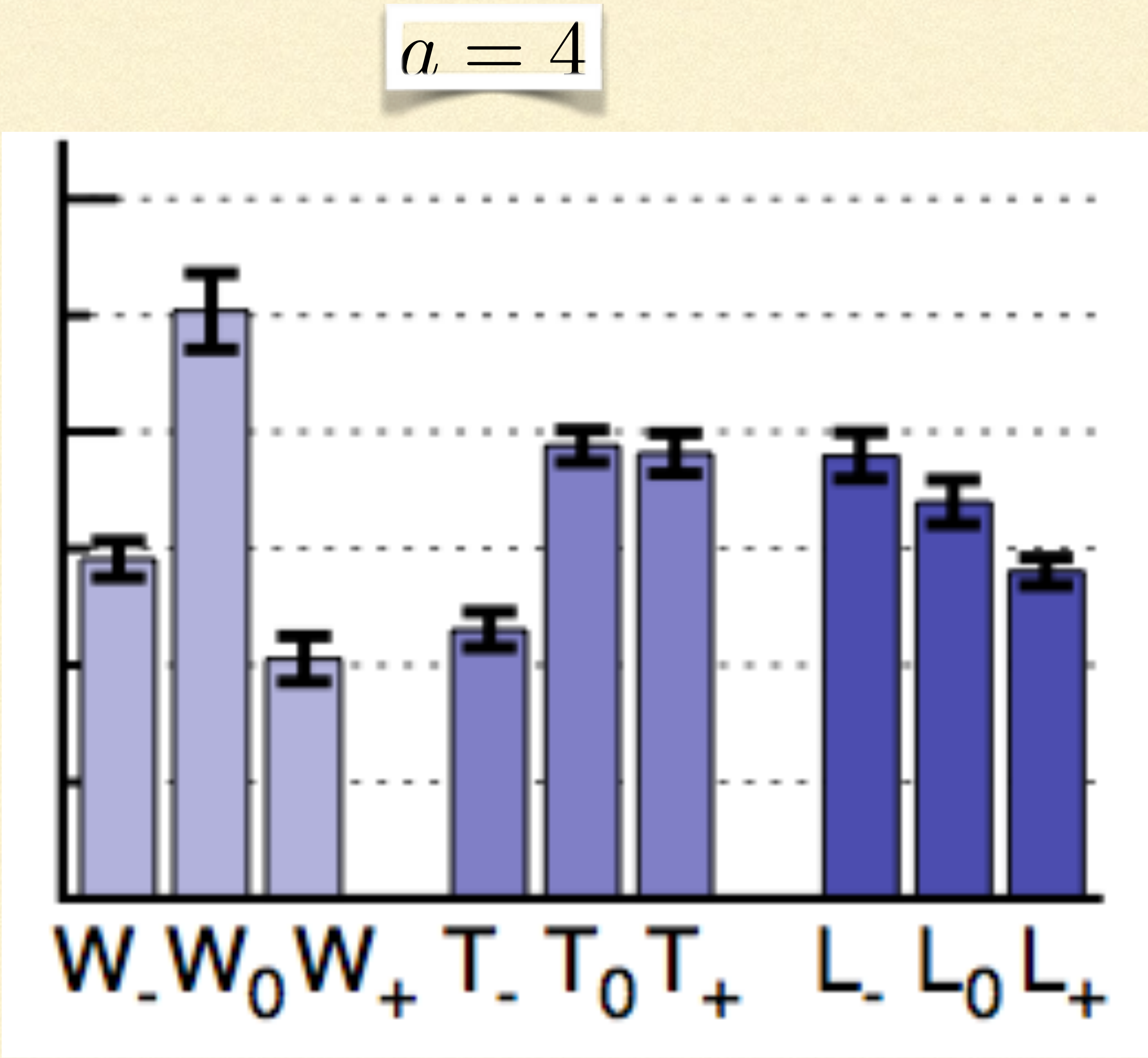
Population-level cyclic motions exist and persist (about 1 turn in 35 game rounds). Cycling direction is counter-clockwise.



WHY SOCIAL-LEVEL CYCLING?

- Cannot be explained by Nash equilibrium theory (infinite rationality).
- Cannot be explained by assuming players make decisions independently of each other.
- Let's get inspirations from empirical data!!!

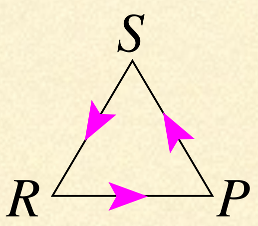
CONDITIONAL RESPONSES



- play outcome:

W (win), T (tie), L (lose)

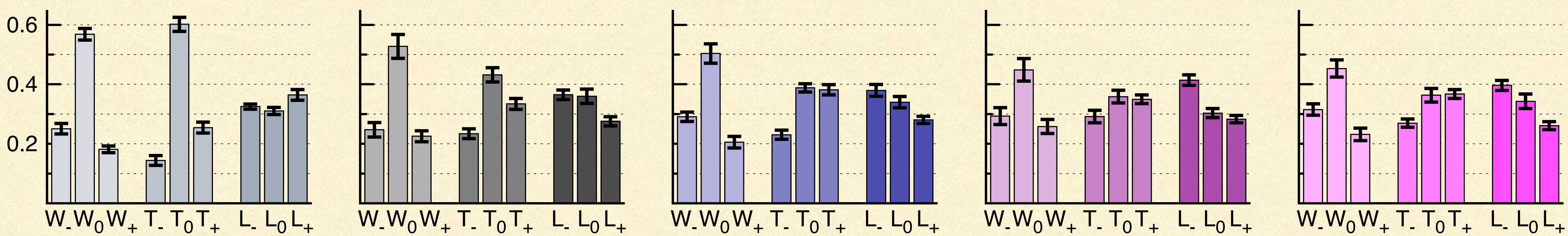
- e.g., if W (win) , next step:

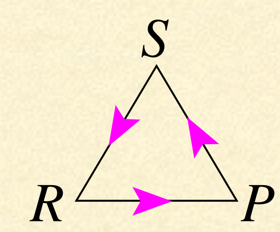


keep old action (prob W_0)
shift action clockwise (prob W_-)
shift action counter-clockwise (prob W_+)

CONDITIONAL RESPONSES

$a = 1.1$ $a = 2$ $a = 4$ $a = 9$ $a = 100$





MODEL BASED ON CR-STRATEGY

At each step, every player of the population chooses an action in a probabilistic way of conditional response:

Given the outcome of the current play being

$$O \in \{W, T, L\}$$

In the next play, the player will choose to

—keep the same action with probability O_0

—shift action clock wisely (R->S, S->R, P->R) with probability O_-

—shift action counter clock wisely (R->P, P->S, S->R) with probability O_+

$$M_{cr}[\mathbf{s}'|\mathbf{s}] = \sum_{n_{rr},n_{pp},\ldots,n_{sr}} \frac{n_R!\; n_P!\; n_S!\; \delta^{n_R}_{2n_{rr}+n_{sr}+n_{rp}}\; \delta^{n_P}_{2n_{pp}+n_{rp}+n_{ps}}\; \delta^{n_S}_{2n_{ss}+n_{ps}+n_{sr}}}{(N-1)!!\; 2^{n_{rr}} n_{rr}!\; 2^{n_{pp}} n_{pp}!\; 2^{n_{ss}} n_{ss}!\; n_{rp}!\; n_{ps}!\; n_{sr}!}$$

$$\times \sum_{n_{rr}^{rr},\ldots,n_{rr}^{sr}} \frac{n_{rr}!\; T_0^{2n_{rr}^{rr}}\; T_+^{2n_{rr}^{pp}}\; T_-^{2n_{rr}^{ss}}\; (2T_+T_0)^{n_{rr}^{rp}}\; (2T_+T_-)^{n_{rr}^{ps}}\; (2T_0T_-)^{n_{rr}^{sr}}}{n_{rr}^{rr}!\; n_{rr}^{pp}!\; n_{rr}^{ss}!\; n_{rr}^{rp}!\; n_{rr}^{ps}!\; n_{rr}^{sr}!} \delta^{n_{rr}}_{n_{rr}^{rr}+\ldots+n_{rr}^{sr}}$$

$$\times \sum_{n_{pp}^{rr},\ldots,n_{pp}^{sr}} \frac{n_{pp}!\; T_-^{2n_{pp}^{rr}}\; T_0^{2n_{pp}^{pp}}\; T_+^{2n_{pp}^{ss}}\; (2T_0T_-)^{n_{pp}^{rp}}\; (2T_+T_0)^{n_{pp}^{ps}}\; (2T_+T_-)^{n_{pp}^{sr}}}{n_{pp}^{rr}!\; n_{pp}^{pp}!\; n_{pp}^{ss}!\; n_{pp}^{rp}!\; n_{pp}^{ps}!\; n_{pp}^{sr}!} \delta^{n_{pp}}_{n_{pp}^{rr}+\ldots+n_{pp}^{sr}}$$

$$\times \sum_{n_{ss}^{rr},\ldots,n_{ss}^{sr}} \frac{n_{ss}!\; T_+^{2n_{ss}^{rr}}\; T_-^{2n_{ss}^{pp}}\; T_0^{2n_{ss}^{ss}}\; (2T_+T_-)^{n_{ss}^{rp}}\; (2T_0T_-)^{n_{ss}^{ps}}\; (2T_+T_0)^{n_{ss}^{sr}}}{n_{ss}^{rr}!\; n_{ss}^{pp}!\; n_{ss}^{ss}!\; n_{ss}^{rp}!\; n_{ss}^{ps}!\; n_{ss}^{sr}!} \delta^{n_{ss}}_{n_{ss}^{rr}+\ldots+n_{ss}^{sr}}$$

$$\times \sum_{n_{rp}^{rr},\ldots,n_{rp}^{sr}} \frac{n_{rp}!\; \delta^{n_{rp}}_{n_{rp}^{rr}+\ldots+n_{rp}^{sr}}}{n_{rp}^{rr}!\; n_{rp}^{pp}!\; n_{rp}^{ss}!\; n_{rp}^{rp}!\; n_{rp}^{ps}!\; n_{rp}^{sr}!} (W_-L_0)^{n_{rp}^{rr}}\; (W_0L_+)^{n_{rp}^{pp}}\; (W_+L_-)^{n_{rp}^{ss}}$$

$$\times\; (W_0L_0 + W_-L_+)^{n_{rp}^{rp}}\; (W_+L_+ + W_0L_-)^{n_{rp}^{ps}}\; (W_+L_0 + W_-L_-)^{n_{rp}^{sr}}$$

$$\times \sum_{n_{ps}^{rr},\ldots,n_{ps}^{sr}} \frac{n_{ps}!\; \delta^{n_{ps}}_{n_{ps}^{rr}+\ldots+n_{ps}^{sr}}}{n_{ps}^{rr}!\; n_{ps}^{pp}!\; n_{ps}^{ss}!\; n_{ps}^{rp}!\; n_{ps}^{ps}!\; n_{ps}^{sr}!} (W_+L_-)^{n_{ps}^{rr}}\; (W_-L_0)^{n_{ps}^{pp}}\; (W_0L_+)^{n_{ps}^{ss}}$$

$$\times\; (W_+L_0 + W_-L_-)^{n_{ps}^{rp}}\; (W_0L_0 + W_-L_+)^{n_{ps}^{ps}}\; (W_+L_+ + W_0L_-)^{n_{ps}^{sr}}$$

$$\times \sum_{n_{sr}^{rr},\ldots,n_{sr}^{sr}} \frac{n_{sr}!\; \delta^{n_{ps}}_{n_{sr}^{rr}+\ldots+n_{sr}^{sr}}}{n_{sr}^{rr}!\; n_{sr}^{pp}!\; n_{sr}^{ss}!\; n_{sr}^{rp}!\; n_{sr}^{ps}!\; n_{sr}^{sr}!} (W_0L_+)^{n_{sr}^{rr}}\; (W_+L_-)^{n_{sr}^{pp}}\; (W_-L_0)^{n_{sr}^{ss}}$$

$$\times\; (W_+L_+ + W_0L_-)^{n_{sr}^{rp}}\; (W_+L_0 + W_-L_-)^{n_{sr}^{ps}}\; (W_0L_0 + W_-L_+)^{n_{sr}^{sr}}$$

$$\times\; \delta^{n'_R}_{2(n_{rr}^{rr}+n_{pp}^{rr}+n_{ss}^{rr}+n_{rp}^{rr}+n_{ps}^{rr}+n_{sr}^{rr})+(n_{rr}^{sr}+n_{pp}^{sr}+n_{ss}^{sr}+n_{rp}^{sr}+n_{ps}^{sr}+n_{sr}^{sr})+(n_{rr}^{rp}+n_{pp}^{rp}+n_{ss}^{rp}+n_{rp}^{rp}+n_{ps}^{rp}+n_{sr}^{rp})}$$

$$\times\; \delta^{n'_P}_{2(n_{rr}^{pp}+n_{pp}^{pp}+n_{ss}^{pp}+n_{rp}^{pp}+n_{ps}^{pp}+n_{sr}^{pp})+(n_{rr}^{rp}+n_{pp}^{rp}+n_{ss}^{rp}+n_{rp}^{rp}+n_{ps}^{rp}+n_{sr}^{rp})+(n_{rr}^{ps}+n_{pp}^{ps}+n_{ss}^{ps}+n_{rp}^{ps}+n_{ps}^{ps}+n_{sr}^{ps})}$$

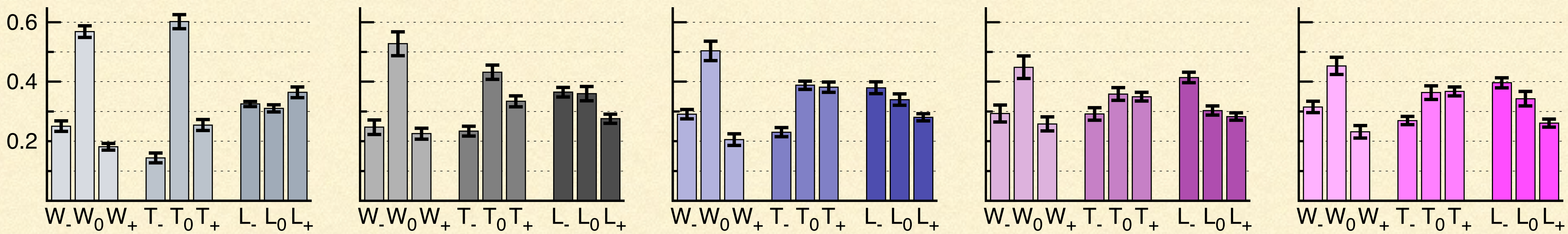
$$\times\; \delta^{n'_S}_{2(n_{rr}^{ss}+n_{pp}^{ss}+n_{ss}^{ss}+n_{rp}^{ss}+n_{ps}^{ss}+n_{sr}^{ss})+(n_{rr}^{ps}+n_{pp}^{ps}+n_{ss}^{ps}+n_{rp}^{ps}+n_{ps}^{ps}+n_{sr}^{ps})+(n_{rr}^{sr}+n_{pp}^{sr}+n_{ss}^{sr}+n_{rp}^{sr}+n_{ps}^{sr}+n_{sr}^{sr})}$$

$a = 1.1$	$a = 2$	$a = 4$	$a = 9$	$a = 100$

$f_{1,300}$

mean	0.031	0.027	0.031	0.022	0.018
s.d.	0.019	0.029	0.026	0.027	0.025
s.e.m.	0.006	0.008	0.008	0.008	0.007

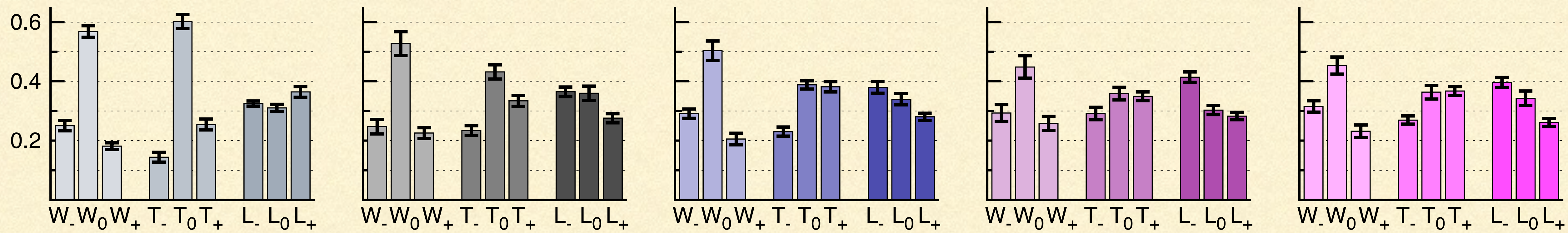
	$a = 1.1$	$a = 2$	$a = 4$	$a = 9$	$a = 100$
			$f_{1,300}$		
mean	0.031	0.027	0.031	0.022	0.018
s.d.	0.019	0.029	0.026	0.027	0.025
s.e.m.	0.006	0.008	0.008	0.008	0.007



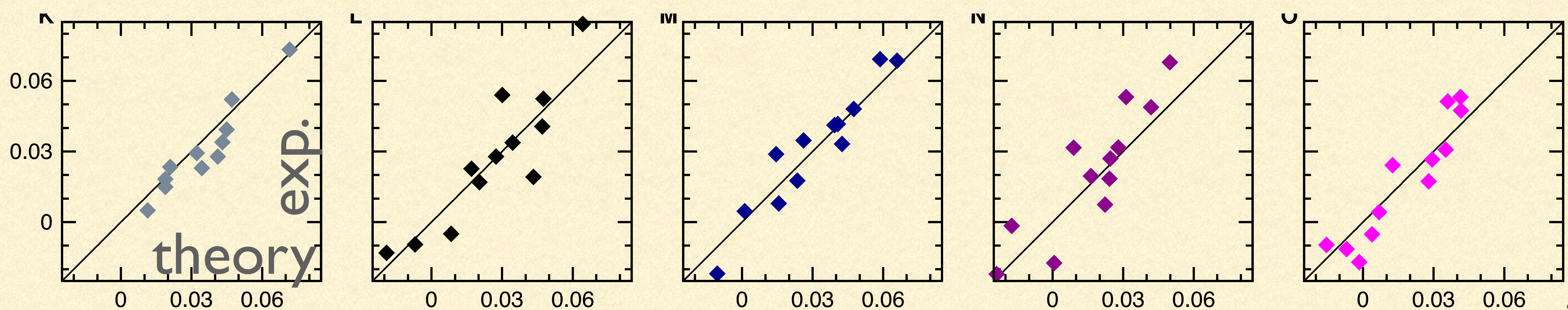
model:	0.035	0.026	0.030	0.018	0.017

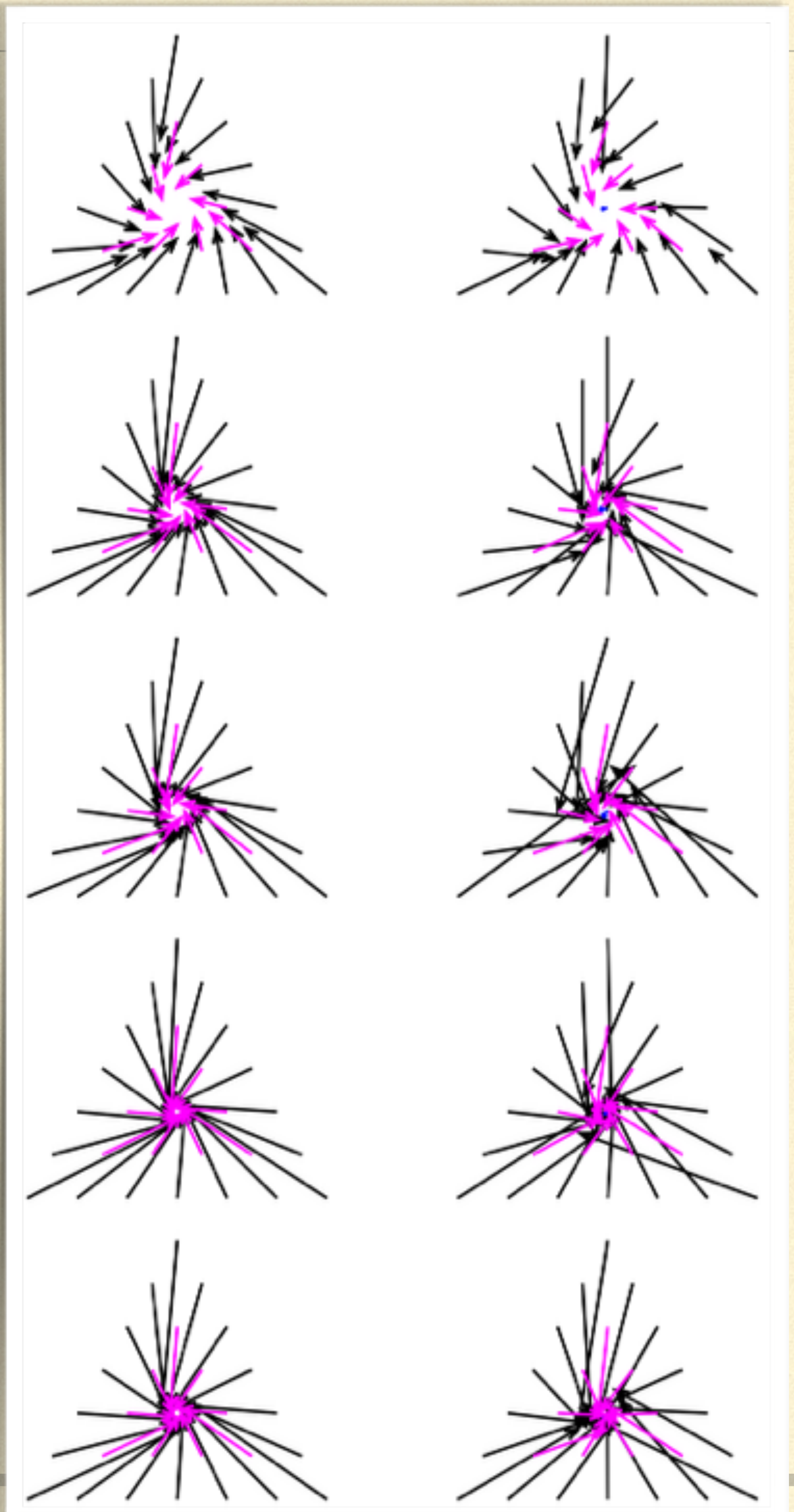
$f_{1,300}$

	$a = 1.1$	$a = 2$	$a = 4$	$a = 9$	$a = 100$
mean	0.031	0.027	0.031	0.022	0.018
s.d.	0.019	0.029	0.026	0.027	0.025
s.e.m.	0.006	0.008	0.008	0.008	0.007



model: 0.035 0.026 0.030 0.018 0.017

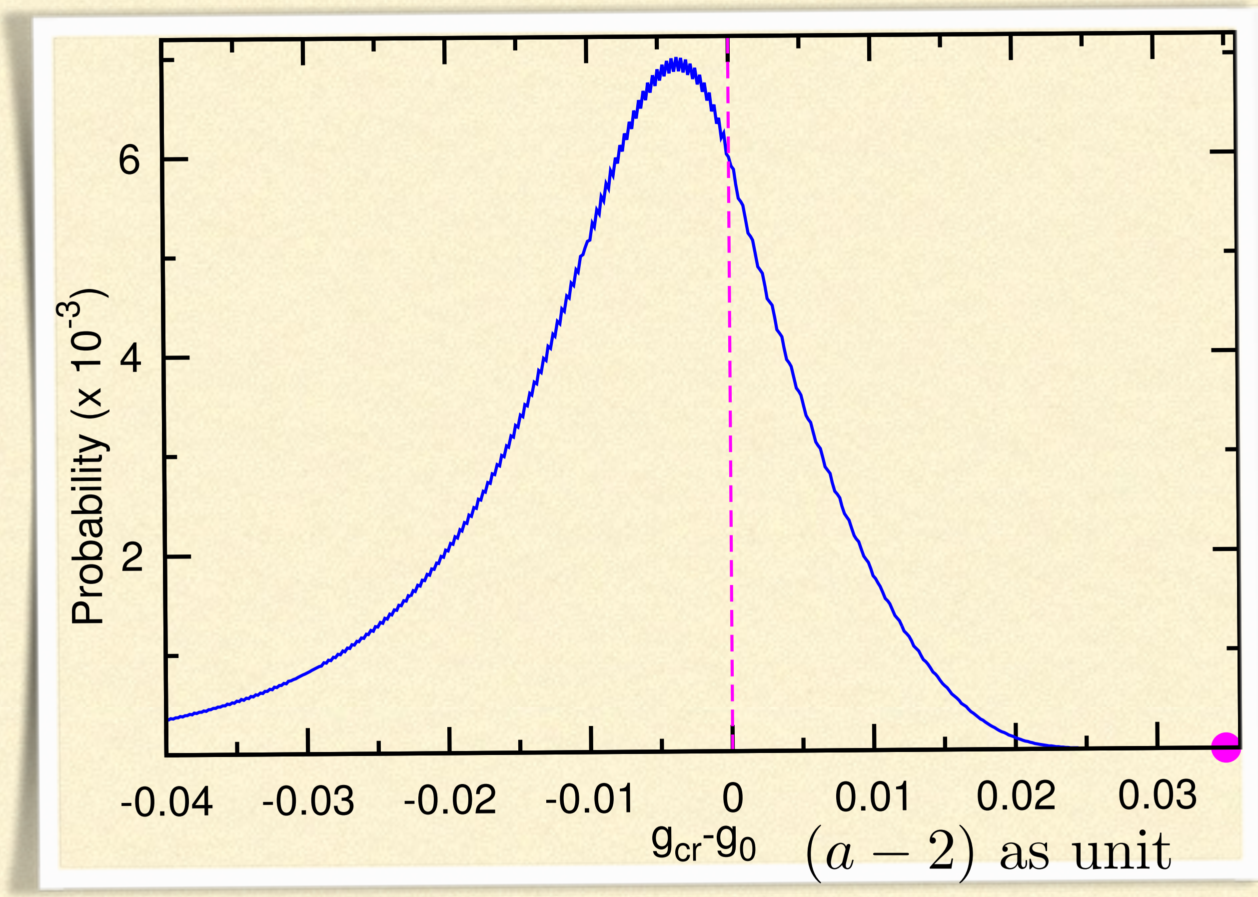




CR-MODEL EXPLAINS CYCLING

- social cycling can be quantitatively explained by the model of conditional response.
- If a player wins over her opponent her opponent in one play, her probability of repeating the same action is considerably higher than her probabilities of shifting actions.
- If a player loses to her opponent in one play, she is more likely to shift action clockwise (R->S, P->R, S->P) than either to keep the old action or to shift action counter-clockwise.

BENEFIT OF CR-STRATEGY?



$g_0 = \frac{1+a}{3}$ NE mixed-strategy

g_{cr} CR-strategy

- 2,400,000,000 CR-strategies sampled uniformly at random to obtain the mean payoff distribution.
- CR-strategy has high probability of being inferior to NE mixed-strategy.
- Yet, optimized CR-strategies can outperform NE mixed-strategy by10% (for population size N=6).
- Empirical mean payoff slightly outperforms NE mixed-strategy.

SOME GOOD CR-STRATEGIES (1)

$$W_- = 0.002 \quad T_- = 0.067 \quad L_- = 0.003$$
$$W_0 = 0.998 \quad T_0 = 0.823 \quad L_0 = 0.994$$
$$W_+ = 0.000 \quad T_+ = 0.110 \quad L_+ = 0.003$$

lazy, but not too lazy

$$f_{cr} = 0.003 \quad g_{cr} = g_0 + 0.035(a-2)$$

SOME GOOD CR-STRATEGIES (2)

$$W_- = 0.995 \quad T_- = 0.800 \quad L_- = 0.988$$
$$W_0 = 0.004 \quad T_0 = 0.142 \quad L_0 = 0.000$$
$$W_+ = 0.001 \quad T_+ = 0.058 \quad L_+ = 0.012$$

coordinated

$$f_{cr} = -0.190 \quad g_{cr} = g_0 + 0.034(a-2)$$

SOME GOOD CR-STRATEGIES (3)

$$W_- = 0.001 \quad T_- = 0.063 \quad L_- = 0.989$$
$$W_0 = 0.994 \quad T_0 = 0.146 \quad L_0 = 0.010$$
$$W_+ = 0.005 \quad T_+ = 0.791 \quad L_+ = 0.001$$

win-stay, lose-shift

$$f_{cr} = 0.189 \quad g_{cr} = g_0 + 0.033(a-2)$$

OUTLOOK

- CR-strategy leads to social cycling, and may even lead to social efficiency. Yet how to **optimize** its parameters **by learning**?

- **Whether Conditional response is a basic decision-making mechanism of the human brain or just a consequence of more fundamental neural mechanisms is a challenging issue for future studies.**

- Let data speak.

AKNOWLEDGEMENTS

- **王志坚、许彬，参与实验的360名浙江大学同学，和实验员**
- 欧阳钟灿老师
- 理论物理国家重点实验室，国家自然科学基金委，中科院交叉基金
- 理论物理研究所hpc计算机集群，及金洪波博士

数据

懂数据

超越数据

（但别超太多）