Cooperation in Social Dilemmas

Christoph Hauert
Department of Mathematics, UBC Vancouver, Canada

www.evoludo.org

Sanctioning institutions for governing the commons

Public goods games

Games and population dynamics

- Large population, players interact in randomly formed groups of size N.
- Two strategic types
  - cooperators x - contribute to common pool at cost c.
  - defectors y - contribute nothing
- Total contributions are multiplied by $r > 1$ and equally split among all *other* participants (irrespective of their type):

$$P_y = \frac{rc}{N-1}x(N-1) = rc\,x$$

x: frequency of cooperators

$$P_x = P_y - c \qquad = (rx-1)c$$

- Payoffs translate into reproductive fitness.

$$\dot{x} = x(P_x - \bar{P}) = x(1-x)(P_x - P_y)$$

↳ Cooperators go extinct.

Punishment

Promoting cooperation - part I

- Punishment is costly - punisher pays γ, punishment fine β.
- Three strategic types
  - cooperators x - contribute to public goods, do not punish
  - defectors y - do not contribute, do not punish
  - peer punishers w - contribute and punish those that did not
- Payoffs

$$\begin{aligned} P_x &= \underbrace{B - c}_{\text{public goods}} \\ P_y &= B \quad -(N-1)w\,\beta \\ P_w &= B - c \quad \underbrace{-(N-1)y\,\gamma}_{\text{punishment}} \end{aligned}$$

$$B = rc(x + w)$$ benefits from public good

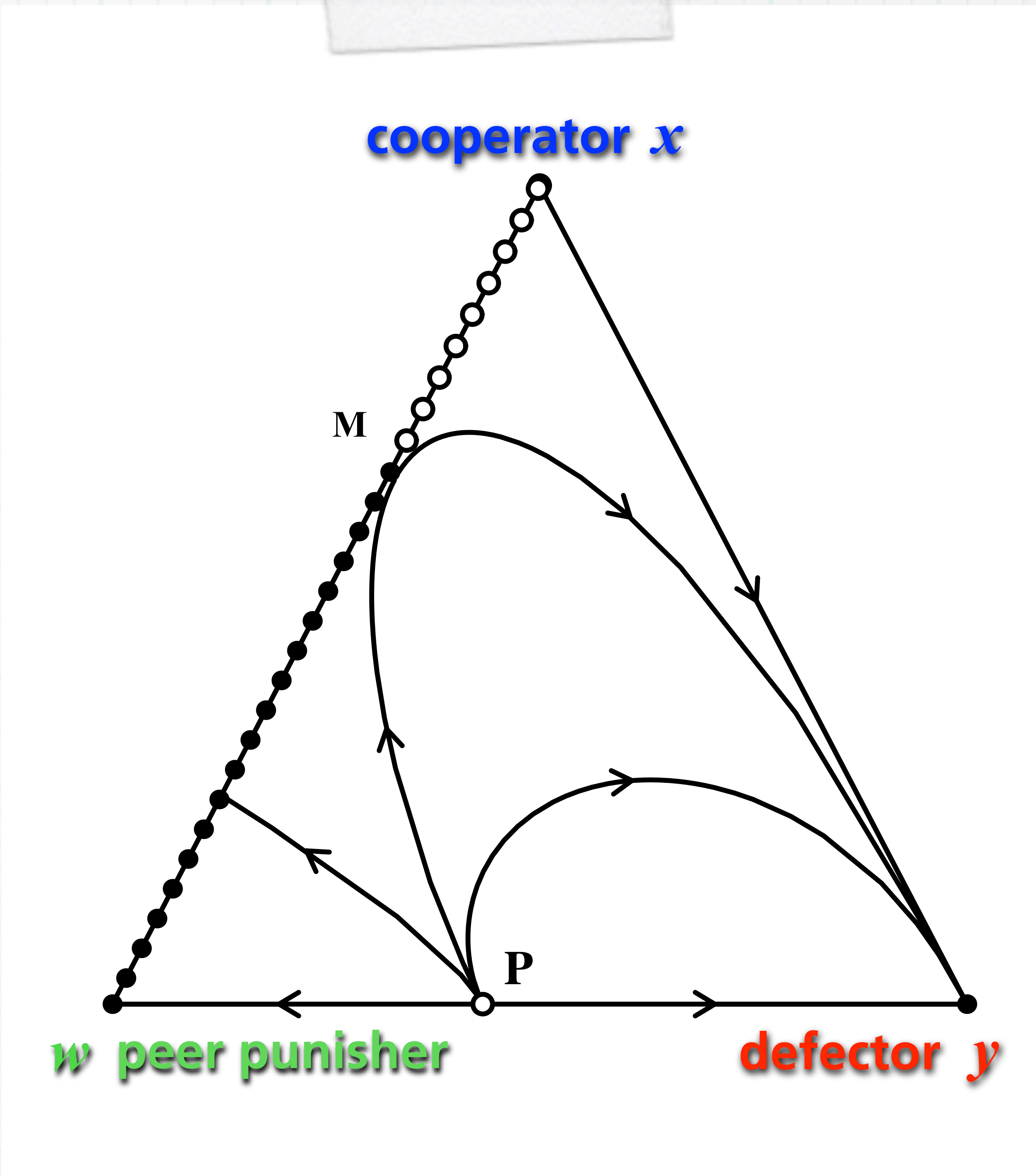
Punishment

Effects of punishment

- Defection is the only stable state.
- No selection in populations of cooperators and peer punishers (line of fixed points).

↳ Cooperators pave the way for defectors.

↳ How can punishment gain a foothold in the population?



*Sigmund, Hauert & Nowak (2001) PNAS **98** 10757.*

Volunteering

Promoting cooperation - part II

- Participation in public goods interactions is voluntary.
- Joint effort is risky - potential for high costs and large benefits.
- Risk averse individuals obtain small but fixed payoff σ

  ($0 < \sigma < (r - 1)c$, better than mutual defection but worse than mutual cooperation).
- Three strategic types
  - cooperators x - contribute to public goods
  - defectors y - do not contribute
  - loners z - refuse to participate

↳ Single participant receives σ.

collective hunt

raiding

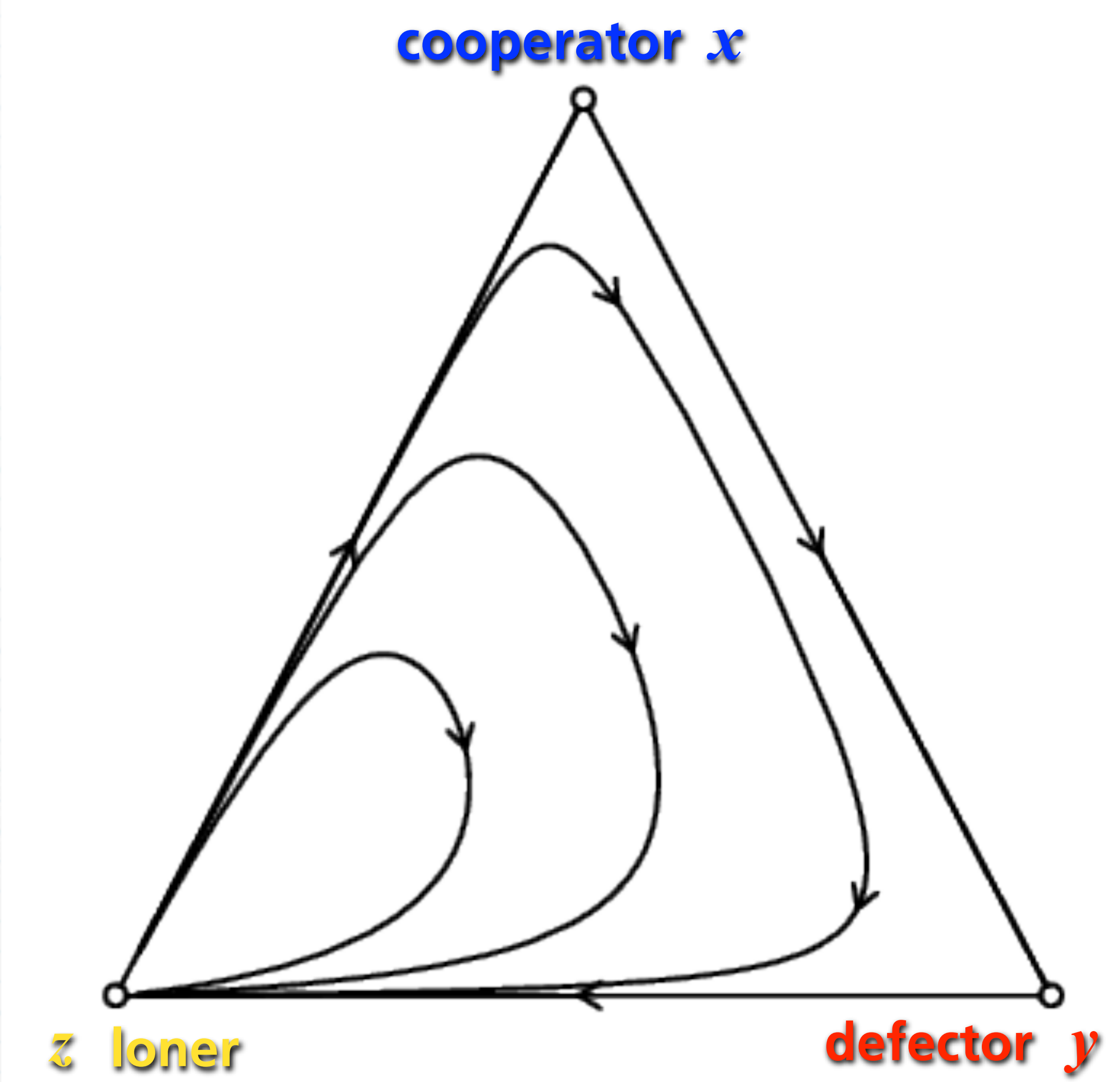
Volunteering

Theory

- Payoffs

$$
\begin{aligned}
P_y &= \sigma z^{N-1} + (1 - z^{N-1}) rc \tfrac{x}{1-z} \\
P_x &= P_y - (1 - z^{N-1}) c \\
P_z &= \sigma
\end{aligned}
$$

- Rock-Scissors-Paper type cyclic dominance along boundary of S_3.
- Loners provide an escape hatch out of states of mutual defection - but this is a fleeting state.



*Hauert, De Monte, Hofbauer & Sigmund (2002) Science **296** 1129.*

Volunteering & Punishment

Promoting cooperation - part III

- Four strategic types
  - cooperators x - contribute to public goods, do not punish
  - defectors y - participate but do not contribute, do not punish
  - loners z - do not participate
  - peer punishers w - contribute and punish
- Allow for second order punishment - punish those that failed to punish (α controls strength).
- Payoffs

$$\begin{aligned}
P_x &= \underbrace{P'_x}_{\text{voluntary public goods}} \; \underbrace{-\alpha\beta w(N-1)(1-(1-y)^{N-2})}_{\text{punishment}} \\
P_y &= P'_y \; -\beta w(N-1) \\
P_z &= \sigma \\
P_w &= P'_x \; -\alpha\gamma x(N-1)(1-(1-y)^{N-2}) \\
& \quad -\gamma y(N-1)
\end{aligned}$$

voluntary public goods

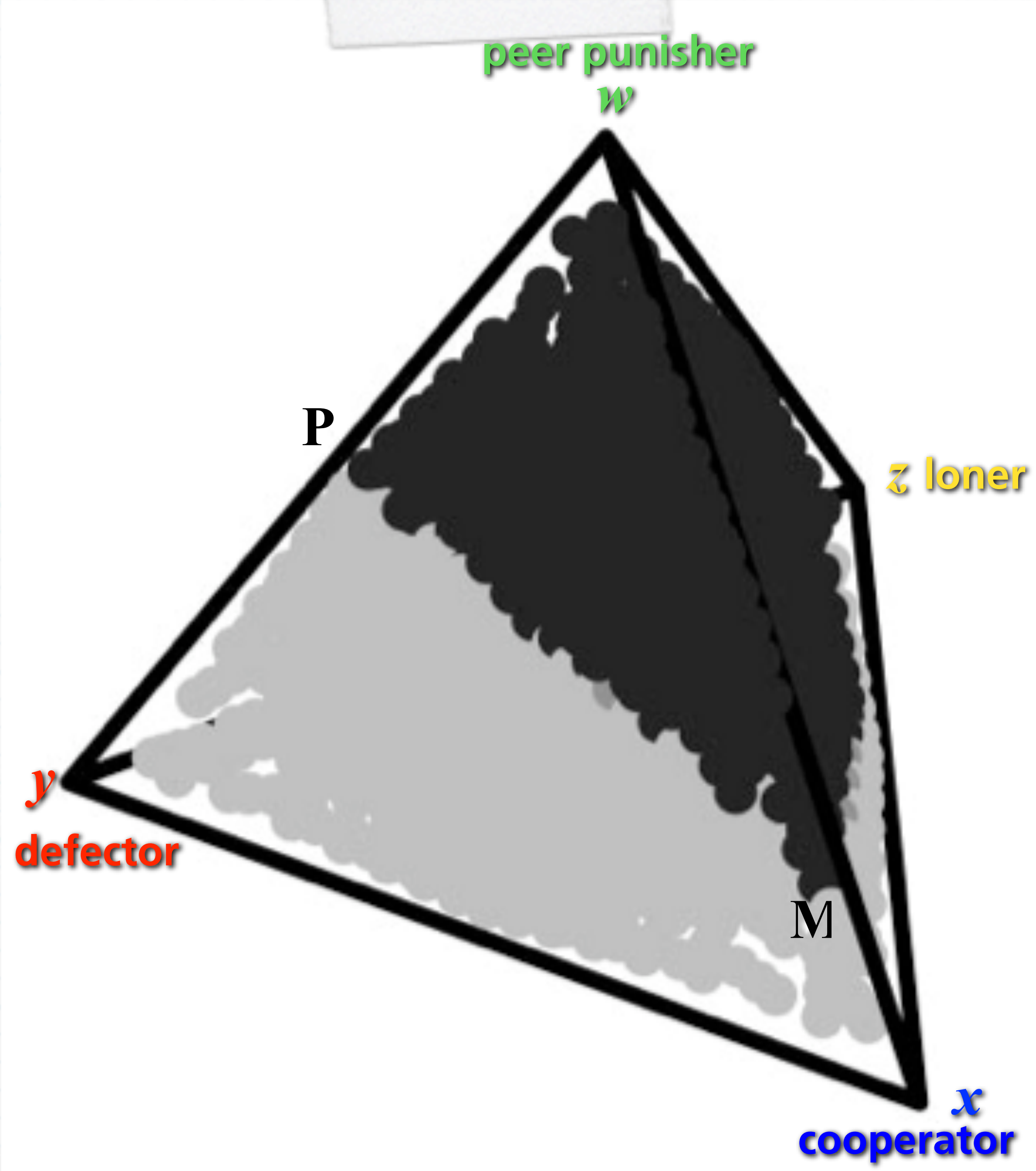
punishment

Volunteering & Punishment

Population dynamics

- Replicator dynamics exhibits two basins of attraction:
  - neutral mixtures of punishers and cooperators (line of fixed points).
  - loners only.
- Fails to explain the evolution of punishment.
- Second order punishment barely affects the dynamics.
- Degenerate dynamics - long term outcome unclear.
- Stochastic model.

Brandt, Hauert, Sigmund (2006)
*PNAS **103** 495.*



Finite Populations

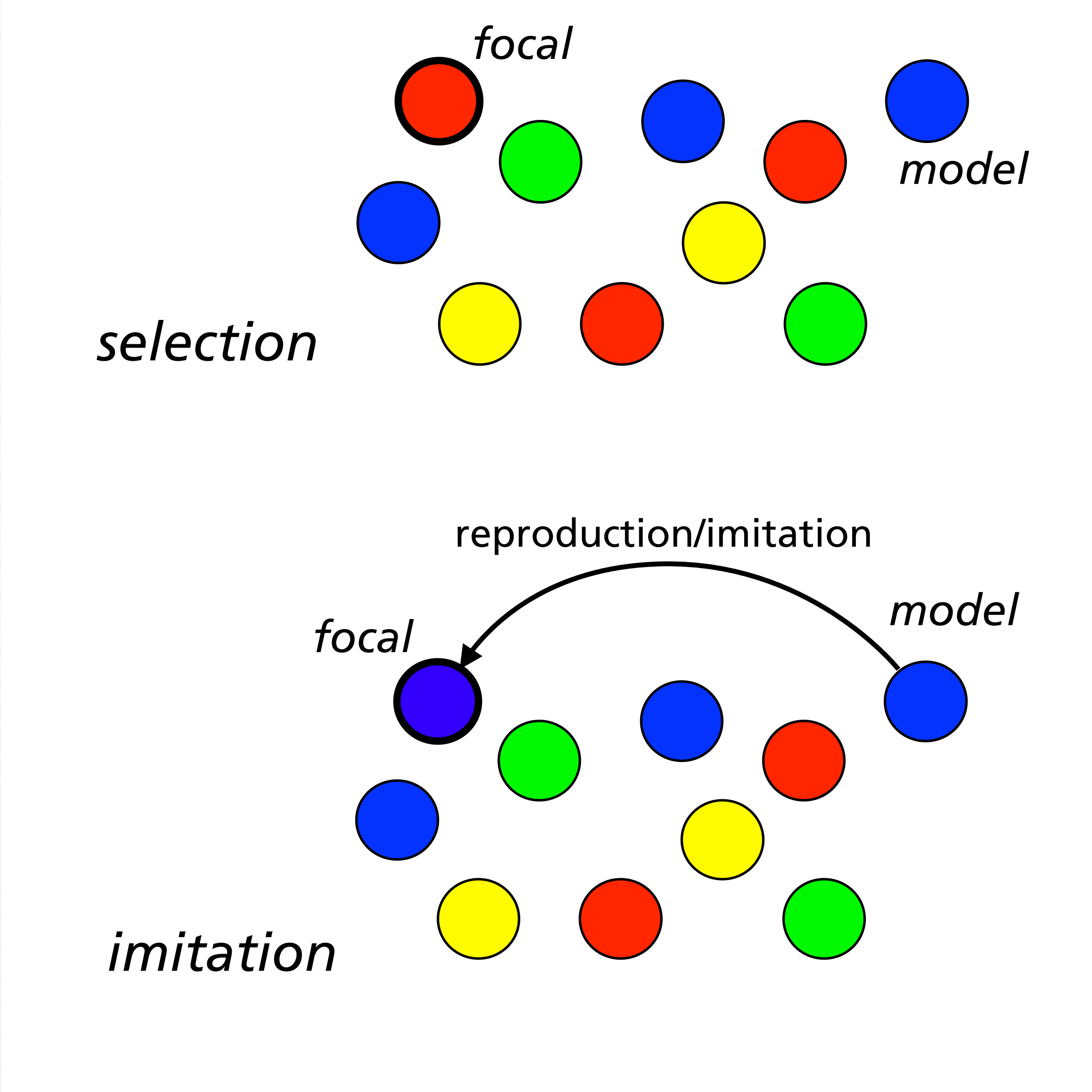
Genetic reproduction or social imitation

- Interaction:
  - Random sampling of interaction group without replacement.
- Evolution:
  - Randomly choose focal individual i.
  - Randomly choose model individual j.
  - Focal individual adopts strategy of model with probability proportional to payoff difference $P_j - P_i$:

$$\frac{1}{1+\exp\left[-s(P_j - P_i)\right]}$$

$s \geq 0$: strength of selection

- $s \to 0$: random selection
- $s \to +\infty$: deterministic selection



Finite populations

Stochastic dynamics

- Rare mutations μ:
  - population is homogeneous most of the time.
  - occasionally a single mutant strategy occurs.
  - mutant disappears or takes over the entire population before next mutation occurs.

↳ Stochastic dynamics along edges of simplex S_n.

- Probability that type i increases competing against type j, T^+_{ij}:

$$T^+_{ij} = \frac{X_i}{M}\frac{M - X_i}{M}\frac{1}{1 + \exp\left[-s(P_i - P_j)\right]}$$

M: population size, X_i: number of type i individuals, M-X_i: number of j types.

- Fixation probability of single i mutant in j population.

$$\rho_{ij} = \frac{1}{\sum_{k=0}^{M-1}\prod_{X_i=1}^{k}\frac{T^-_{ij}}{T^+_{ij}}} = \frac{1}{\sum_{k=0}^{M-1}\exp\left[s\sum_{X_i=1}^{k}(P_j - P_i)\right]}$$

↳ Embedded Markov chain for transitions between homogenous states.

Finite populations

Strong imitation $s \to \infty$

- Strong imitation significantly simplifies Markov chain.
  Ex. voluntary public goods games: $\rho_{XY} = \rho_{YZ} = 1$ and $\rho_{ZX} = 1/2$ all other $\rho_{ij} = 0$.

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & 0 & \frac{3}{4} \end{pmatrix} \quad \to \quad \left(\frac{1}{4}, \quad \frac{1}{4}, \quad \frac{1}{2}\right)$$

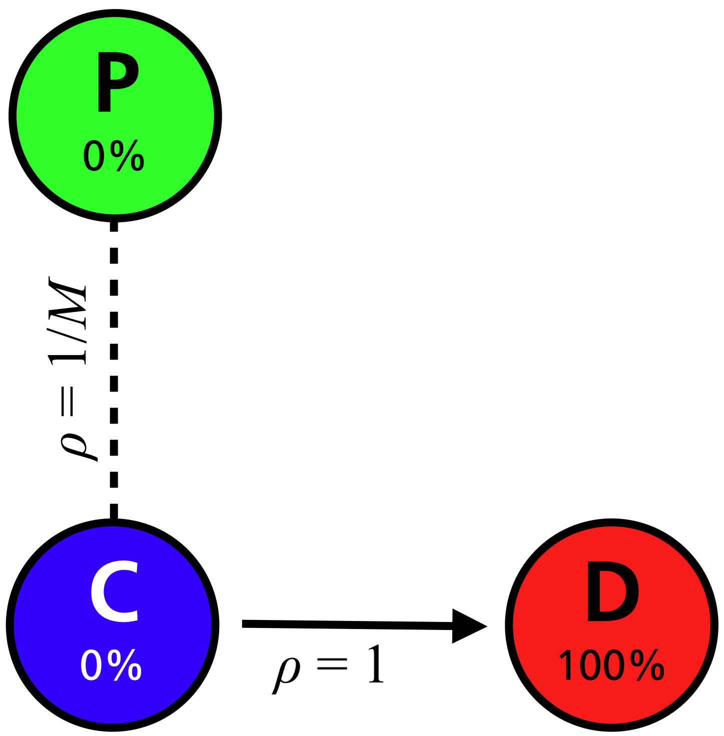
- The result becomes independent of the parameters! (as long as the cyclic dominance $X \to Y \to Z \to X$ persists).

- $\rho_{ZX} = 1/2$ because two cooperators are required to invade a loner population (non-hyperbolic fixed point in replicator equation).

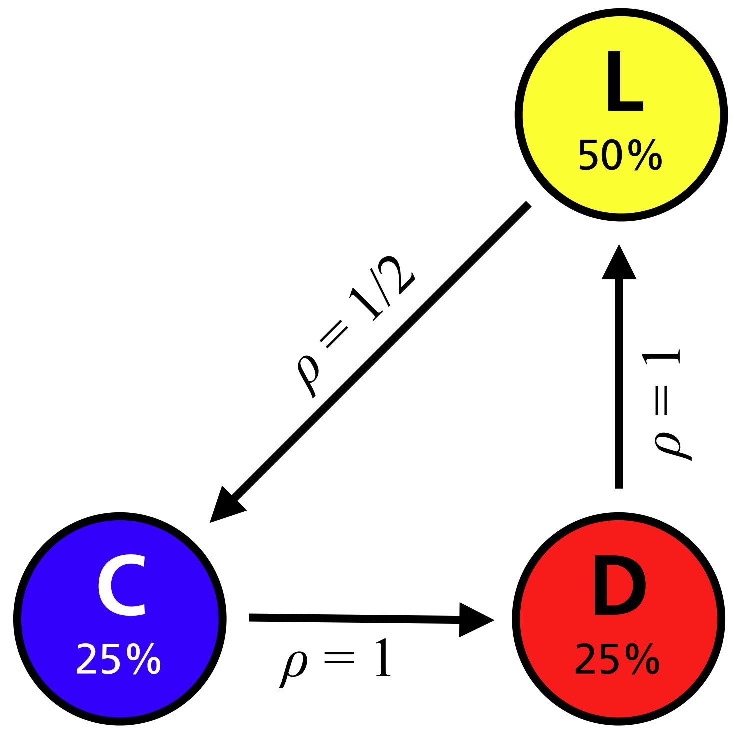
- Neutral evolution (no fitness differences) yields fixation probability of $1/M$ where M denotes the population size.

Volunteering & Punishment

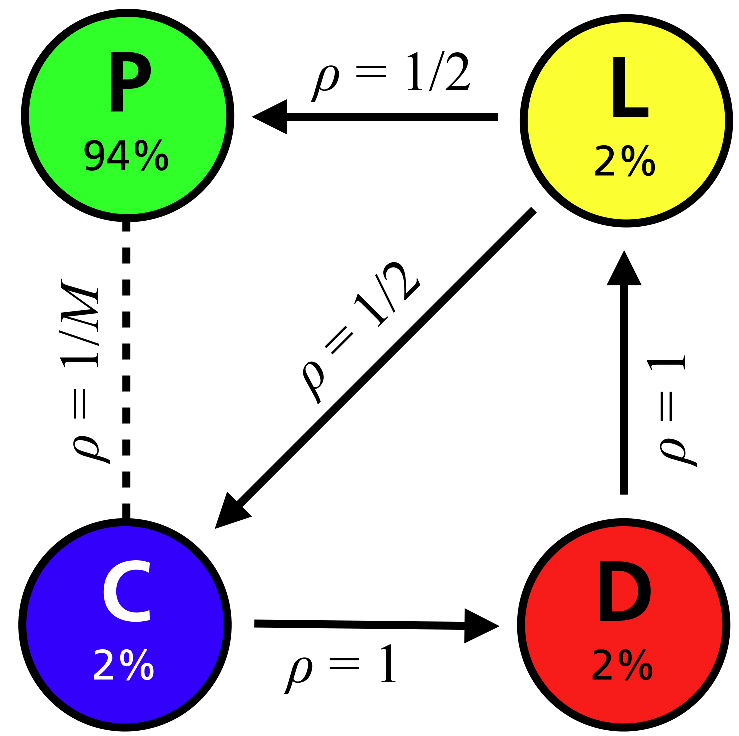
Results



Compulsory public goods games with peer punishment:
defectors rule.



Voluntary public goods games:
cyclic dominance.



Voluntary public goods games with peer punishment:
punishers reign (M=92).

*Hauert, Traulsen, Brandt, Nowak & Sigmund (2007) Science **316** 1905.*

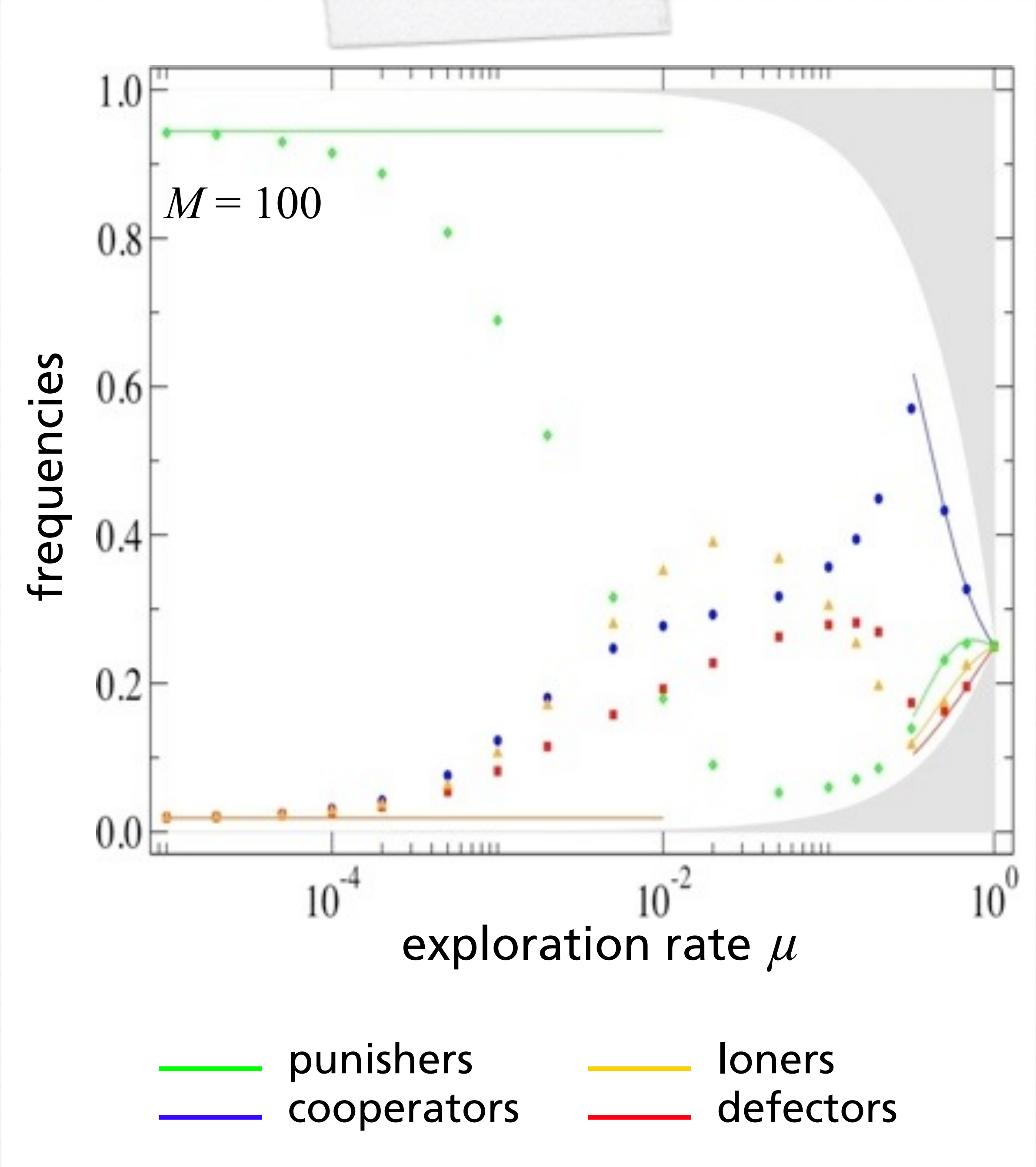
Cultural Evolution

Population dynamics

- In cultural evolution ‘mutation’ rates may not be small - individuals randomly experiment with different strategies.
- For smaller μ, punishment prevails.
  - The stochastic dynamics in finite populations can resolve the problem of establishing altruistic punishment
- Cooperators prevail for large μ.

  (Note that here the contributors also get a return on their own investment - otherwise loners dominate.)
  - Loners are no longer crucial.
  - Punishers are pivotal for the success of mild cooperators (second order free riders).

*Traulsen, Hauert, De Silva, Nowak & Sigmund (2009) PNAS **106** 709-712.*



Sanctioning Institutions

Promoting cooperation - part IV

- Second order free riders (contribute but do not punish) undermine punishing efforts.
  - Even $\alpha=1$ cannot prevent this because cannot identify among contributors.
  - Establish punishment pool where individual contribute *before* engaging in the public goods interaction.
  - Precursor to institutionalized punishment.
  - Easy identification of free riders.
- Five strategic types
  - cooperators x - contribute to public goods, do not punish
  - defectors y - participate but do not contribute, do not punish
  - loners z - do not participate
  - peer punishers w - contribute and punish those that did not contribute
  - pool punishers v - contribute to public goods *and* to punishment pool
- Pool punishers pay an additional amount $G > 0$ into the punishment pool and free riders are fined proportional to the number of pool punishers N_v: $N_v\,G$.

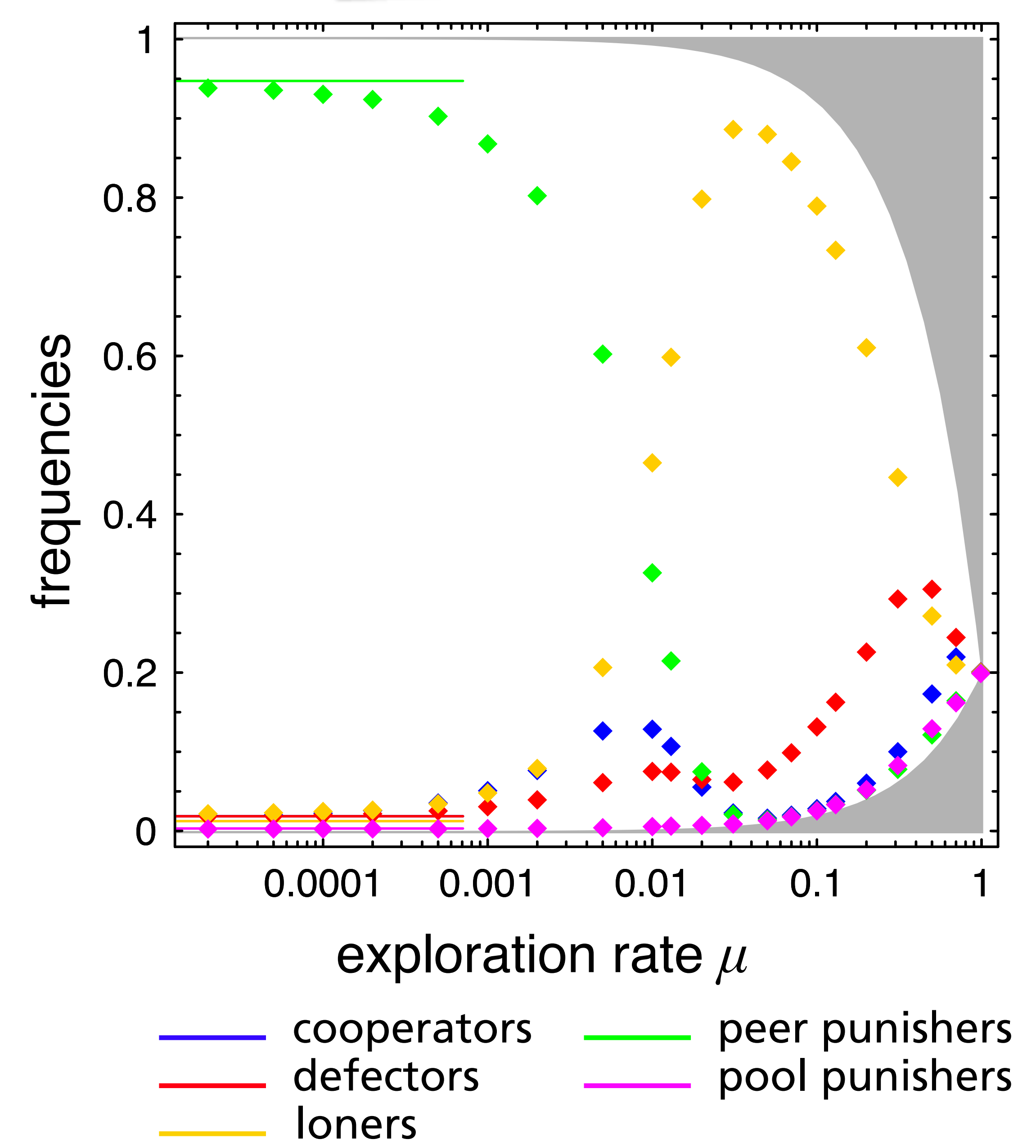
Peer versus pool punishment

No second order punishment, α=0

- Payoffs (infinite populations)

$$
\begin{aligned}
P_x &= \underbrace{P'_x}_{\text{public goods with peer punishment}} && \underbrace{-\alpha G v(N-1)}_{\text{pool punishment}} \\
P_y &= P'_y && -Gv(N-1) \\
P_z &= \sigma && \\
P_w &= P'_w && -\alpha G v(N-1) \\
P_v &= P'_x && -G
\end{aligned}
$$

- Punishment pool only used to punish non-contributors (defectors).
  - For small μ peer-punishment dominates.
  - For large μ the public good collapses.
  - Pool punishment ineffective.
- For compulsory interactions (no loners) defectors dominate.

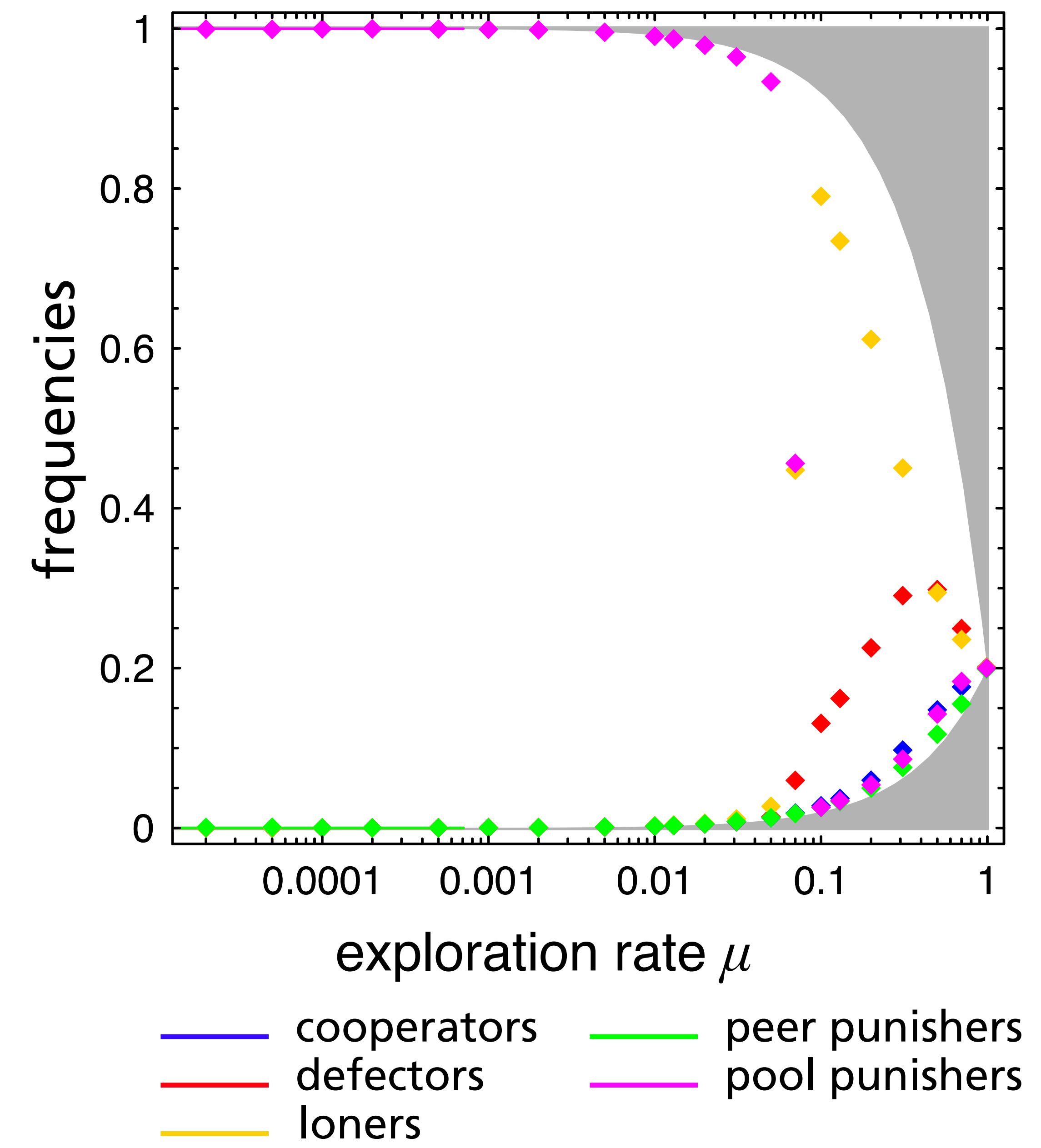


Peer versus pool punishment

With second order punishment, $\alpha=1$

- Punishment pool used to punish non-contributors (defectors) as well as those that do not commit to pool-punishment (cooperators and peer-punishers).
  - Pool punishers prevail for most μ.
  - For very large μ the public good collapses.
- For compulsory interactions (no loners) defectors again dominate.
  - Punishment often fails in compulsory public goods.
  - Preservation of global resources (climate, air, water, fish...).
  - "Mutual coercion mutually [and voluntarily] agreed upon". Hardin, 1968

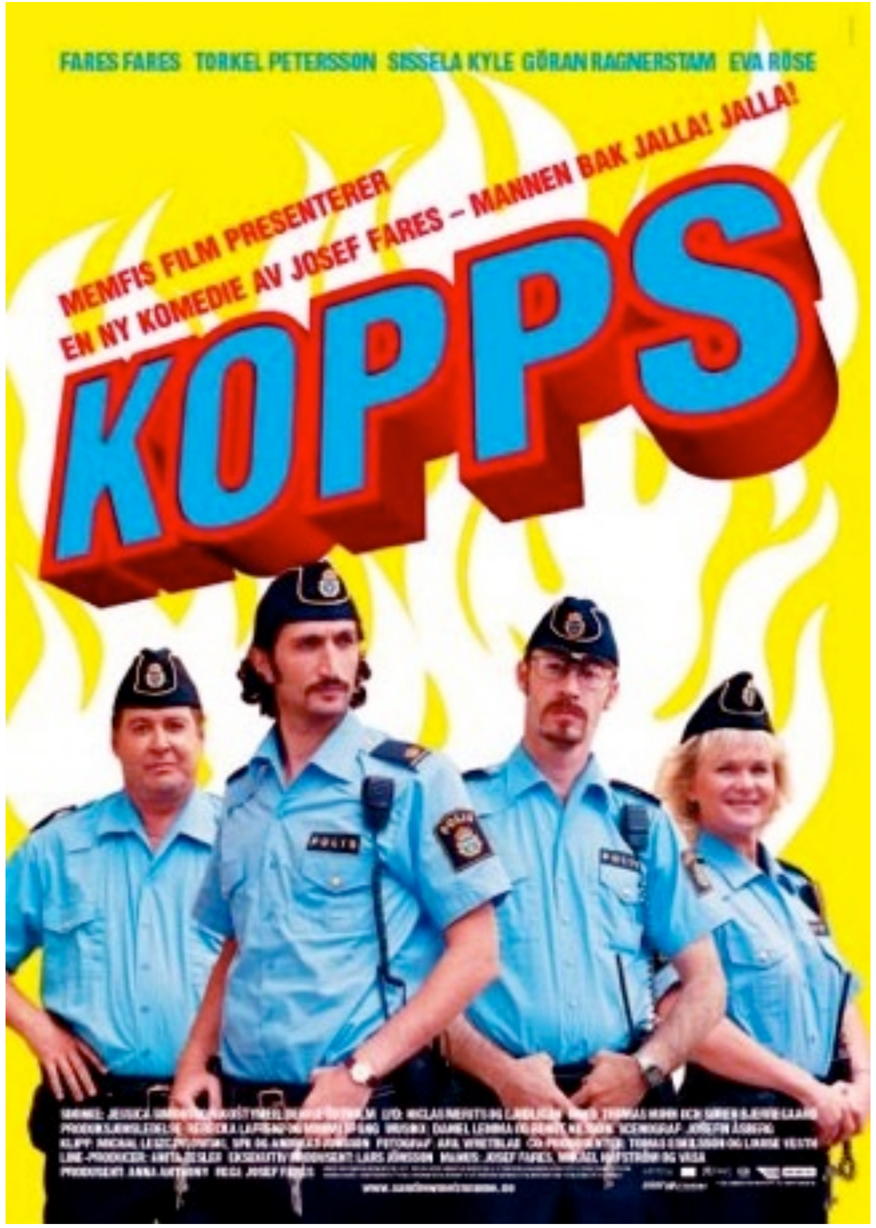
*Sigmund, De Silva, Traulsen & Hauert (2010) Nature **466** 861-863.*



Conclusions

Sanctioning Institutions

- Peer punishment relates to the instinct for revenge.
- Pool punishment as a step towards establishing sanctioning institutions: commit resources to prepare for punishing free-riders.
  - Pool punishment is based on foresight rather than anger.
- Populations of peer punishers are better off than pool punishers.
  - The upkeep of the punishment pool incurs costs.
- With second order punishment, pool punishers prevail.
  - Higher efficiency of peer punishment is traded for greater stability of pool punishment.



Acknowledgements

***Karl Sigmund**, University of Vienna, Austria.*

***Josef Hofbauer**, University of Vienna, Austria.*

***Hannelore De Silva (neé Brandt)**, Vienna University of Economics, Austria.*

***Martin Nowak**, Program for Evolutionary Dynamics, Harvard University.*

***Arne Traulsen**, Max-Planck Institute for Evolutionary Biology, Plön, Germany.*

***Silvia De Monte**, École Normale Superieure, Paris, France.*

VirtualLabs installation on 3D NOVA display
Location: main train station Zürich, Switzerland
Media artist: Chandrasekhar Ramakrishnan

Tutorials:
http://www.univie.ac.at/virtuallabs
and soon http://www.evoludo.org