Prisoner's Dilemma on a Network

Zheng Xia, Tim de Groot, Iga Pregowska, Tristan Farran

January 2026

1 Research questions with quantitative measurements

Let $G = (V, E)$ be a graph with $|V| = N$. Each agent $i \in V$ takes an action $a_i(t) \in \{C, D\}$ at (discrete) time t. Define the cooperation indicator $x_i(t) = I\{a_i(t) = C\}$ and the global cooperation fraction

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t).$$

RQ0 (Baseline existence): Can cooperation persist?

RQ0. For a given network family G, payoff parameters, and update rule, does the system reach *persistent (non-trivial) cooperation* rather than collapse to defection?

Measurements. Choose a burn-in T_{burn} and averaging window W:

$$\bar{C} = \frac{1}{W} \sum_{t=T_{\text{burn}}+1}^{T_{\text{burn}}+W} C(t).$$

If the dynamics admits absorbing states, estimate:

- *Fixation probability:* $\Pr(\text{all-}C)$ and $\Pr(\text{all-}D)$ across runs.
- *Time to absorption/stationarity:* steps until $C(t)$ stabilizes (or absorption occurs).

RQ1 (Network architecture): Which structures promote cooperation?

RQ1. Holding N and mean degree $\langle k \rangle$ fixed, how does network architecture (e.g., lattice, ER, WS, BA, DC-SBM) affect long-run cooperation and local assortment?

Measurements.

- Long-run cooperation level $\bar{C}$.
- *Neighbor cooperation (assortment).* Let $\mathcal{N}(i)$ be neighbors of i and $k_i = |\mathcal{N}(i)|$. Define neighbor mean

$$\bar{x}_{\mathcal{N}(i)}(t) = \frac{1}{k_i} \sum_{j \in \mathcal{N}(i)} x_j(t).$$

  Define Pearson correlation

$$r(t) = \operatorname{Corr}\big(x_i(t),\ \bar{x}_{\mathcal{N}(i)}(t)\big)_{i=1}^{N}.$$

- Report network statistics (as controls/mediators): clustering coefficient, average shortest path length (on the giant component), degree heterogeneity (e.g., $\mathrm{CV}(k)$), modularity (if applicable).

RQ2 (Cooperative clusters): Size, stability, and morphology

RQ2. When cooperation exists, does it form many small islands or large cohesive clusters? How stable are cooperative clusters over time?

Measurements. Let $V_C(t) = \{i \in V : x_i(t) = 1\}$. Define the induced cooperative subgraph $G_C(t) = G[V_C(t)]$. Let connected component sizes in $G_C(t)$ be $\{s_\ell(t)\}_\ell$.

- *Largest cooperative component fraction (order parameter):*

$$S_{\max}(t) = \frac{\max_\ell s_\ell(t)}{N}, \qquad \overline{S_{\max}} = \frac{1}{W} \sum_{t=T_{\text{burn}}+1}^{T_{\text{burn}}+W} S_{\max}(t).$$

- *Cluster size distribution:* histogram of $\{s_\ell(t)\}$ over time and runs.
- *Set stability (Jaccard index):*

$$J(t;\Delta) = \frac{|V_C(t) \cap V_C(t-\Delta)|}{|V_C(t) \cup V_C(t-\Delta)|}.$$

- Or stability of the largest component: overlap of nodes in the largest component across time.

RQ3 (Phase transitions / percolation-like behavior): Are there thresholds?

RQ3. As a control parameter (e.g., temptation T, ratio b/c, noise K, rewiring rate w) varies, is there a threshold at which the cooperative subgraph transitions from fragmented components to a giant cooperative component?

Measurements. Use $S_{\max}(t)$ as an order parameter and define a percolation-style susceptibility from cooperative component counts. Let $n_s(t)$ be the number of cooperative components of size s at time t. Then

$$\chi(t) = \frac{\sum_{s \neq s_{\max}} s^2\, n_s(t)}{\sum_{s \neq s_{\max}} s\, n_s(t)}.$$

Look for sharp changes in $\overline{S_{\max}}$ and peaks in $\bar{\chi}$ (time-averaged $\chi(t)$) across parameter sweeps.

RQ4 (Endogenous rewiring): Does coevolution stabilize cooperation?

RQ4. If edges can rewire (especially with cooperator preference), how does coevolution of network structure and strategies affect cooperation and cluster formation?

Measurements. All above, plus:

- Edge-type mixing fractions: $f_{CC}(t), f_{CD}(t), f_{DD}(t)$.
- Time series of network statistics under rewiring: clustering, modularity, degree distribution moments.
- (Optional) "cascade" sizes from $\Delta C(t)$ or $\Delta S_{\max}(t)$ to probe burstiness.

2 Hypotheses aligned with research questions

H0 (existence)

- **H0.1:** For sufficiently favorable payoff conditions (e.g., lower temptation or higher b/c), and for imitative dynamics, $\bar{C}$ remains above a non-trivial level on structured networks.

H1 (architecture effects at fixed $\langle k \rangle$)

- **H1.1 (clustering helps):** Higher clustering (e.g., lattice / WS at low rewiring) yields larger $\bar{C}$ and higher assortment $r(t)$ than ER networks at the same $\langle k \rangle$.
- **H1.2 (heterogeneity is conditional):** Degree heterogeneity (e.g., BA) can promote or hinder cooperation depending on whether payoffs are accumulated (π_i) or normalized (π_i / k_i) and on the update rule.
- **H1.3 (community structure helps):** DC-SBM with stronger community structure increases $\bar{C}$ and $r(t)$ due to within-community reinforcement.

H2 (cluster morphology)

- **H2.1:** High-cooperation regimes coincide with larger $\overline{S_{\max}}$ and positive assortment $r(t) > 0$.
- **H2.2:** ER-like networks yield more fragmented cooperative patterns (smaller $\overline{S_{\max}}$, lower susceptibility peak) than clustered/modular networks.

H3 (critical transitions)

- **H3.1:** There exists a parameter region where $\overline{S_{\max}}$ rises rapidly and $\bar{\chi}$ peaks, consistent with a percolation-like transition in $G_C(t)$.
- **H3.2:** The apparent threshold shifts with network clustering/modularity (e.g., lower threshold in highly clustered/modular networks).

H4 (rewiring / coevolution)

- **H4:** Cooperator-favoring rewiring increases $f_{CC}(t)$, increases assortment $r(t)$, and raises $\bar{C}$.

3 Implementation plan: networks, controls, and measurements

3.1 Step A: Baseline experimental frame (fair comparisons)

Fix core settings across experiments:

- Network size N (e.g., $500, 1000, 2000$) and target mean degree $\langle k \rangle$ (e.g., 6 or 8).
- Simulation horizon T (e.g., 5,000–20,000), burn-in T_{burn}, and averaging window W.
- Replicates per condition: multiple network realizations and multiple random seeds.

3.2 Step B: Network families to compare (and structural controls)

Recommended minimal set:

1. **2D lattice/grid:** strong spatial structure. Controls: neighborhood type, periodic boundary.
2. **Erdős–Rényi (ER):** $G(N, p)$ with p chosen to match $\langle k \rangle \approx p(N-1)$.
3. **Watts–Strogatz (WS):** ring degree k and rewiring p_{WS} (tunes clustering/path length).
4. **Barabási–Albert (BA):** preferential attachment with parameter m ($\langle k \rangle \approx 2m$).
5. **DC-SBM:** communities. Controls: number of blocks B, $p_{\text{in}}, p_{\text{out}}$ (or mixing μ), and degree correction.

For each generated graph, compute: clustering coefficient, average path length (giant component), modularity, degree heterogeneity (e.g., $\text{CV}(k)$).

3.3 Step C: Interaction model (Prisoner's Dilemma) and update rule

Payoffs. Either use canonical PD payoffs (T, R, P, S) with $T > R > P > S$ (and $2R > T + S$), or the donation game with benefit b and cost c ($b > c > 0$):

$$u(C, C) = b - c, \quad u(C, D) = -c, \quad u(D, C) = b, \quad u(D, D) = 0.$$

Each round, agent i plays with all neighbors and receives accumulated payoff

$$\pi_i(t) = \sum_{j \in \mathcal{N}(i)} u(a_i(t), a_j(t)).$$

Robustness check (important for BA): use normalized payoff $\pi_i(t)/k_i$.

Strategy space (staged).

- Stage 1: unconditional strategies (ALLC, ALLD).
- Stage 2: imitation dynamics (standard for network reciprocity):
  - *Imitate-best:* copy best-performing neighbor.
  - *Fermi rule:* select neighbor j and copy with probability

    $$\Pr(i \leftarrow j) = \frac{1}{1 + \exp(-(\pi_j - \pi_i)/K)},$$

    where $K > 0$ is selection noise ("temperature").

Update schedule. Compare synchronous vs asynchronous (random sequential) updates as a robustness check.

3.4 Step D (optional): Endogenous rewiring module

At each step (or per edge) with rewiring rate w:

1. Select a node i and a neighbor j uniformly at random.
2. If (i, j) is a CD edge, cut it with probability q.
3. Rewire i to a new node k:
   - Uniform selection; or
   - Cooperator-biased: $\Pr(k) \propto 1 + \eta\, I\{a_k = C\}$.

Controls: rewiring rate w, selectivity q, cooperator preference η.

3.5 Step E: Measurements (definition and computation)

At sampled times t:

- Global cooperation $C(t)$ and long-run $\bar{C}$.
- Cooperative subgraph $G_C(t) = G[V_C(t)]$; component sizes $\{s_\ell(t)\}$; order parameter $S_{\max}(t)$ and susceptibility $\chi(t)$.
- Assortment $r(t) = \mathrm{Corr}(x_i(t), \bar{x}_{\mathcal{N}(i)}(t))$.
- Edge-type fractions $f_{CC}(t), f_{CD}(t), f_{DD}(t)$.
- (If rewiring) network statistics over time: clustering, modularity, degree moments.