

SCIENTIFIC REPORTS



OPEN

Direct reciprocity and model-predictive rationality explain network reciprocity over social ties

Fabio Dercole , Fabio Della Rossa  & Carlo Piccardi 

Since M. A. Nowak & R. May's (1992) influential paper, limiting each agent's interactions to a few neighbors in a network of contacts has been proposed as the simplest mechanism to support the evolution of cooperation in biological and socio-economic systems. The network allows cooperative agents to self-assort into clusters, within which they reciprocate cooperation. This (induced) network reciprocity has been observed in several theoretical models and shown to predict the fixation of cooperation under a simple rule: the benefit produced by an act of cooperation must outweigh the cost of cooperating with all neighbors. However, the experimental evidence among humans is controversial: though the rule seems to be confirmed, the underlying modeling assumptions are not. Specifically, models assume that agents update their strategies by imitating better performing neighbors, even though imitation lacks rationality when interactions are far from all-to-all. Indeed, imitation did not emerge in experiments. What did emerge is that humans are conditioned by their own mood and that, when in a cooperative mood, they reciprocate cooperation. To help resolve the controversy, we design a model in which we rationally confront the two main behaviors emerging from experiments—reciprocal cooperation and unconditional defection—in a networked prisoner's dilemma. Rationality is introduced by means of a predictive rule for strategy update and is bounded by the assumed model society. We show that both reciprocity and a multi-step predictive horizon are necessary to stabilize cooperation, and sufficient for its fixation, provided the game benefit-to-cost ratio is larger than a measure of network connectivity. We hence rediscover the rule of network reciprocity, underpinned however by a different evolutionary mechanism.

Cooperation among self-interested agents is a longstanding and still debated puzzle in biology and social sciences, with countless contributions since R. Axelrod, W. D. Hamilton, and R. L. Trivers' seminal works^{1,2} (see ref.³ for a recent review on human cooperation); the topic has also received attention in several fields of engineering⁴.

The standard modeling framework is *evolutionary game theory* (EGT)^{5–7}, in which a game describes the interaction among pairs (or a larger groups⁸) of self-interested agents in a population, a given set of behavioral strategies is confronted, and an evolutionary process links the obtained payoffs to reproduction and death in biology or to strategy update in socio-economic systems. The paradigmatic game used to study the evolution of cooperation is the *prisoner's dilemma* (PD), the two-player-two-option interaction in which a cooperator (option C) provides a benefit b to the opponent at a cost $c < b$ to herself, whereas a defector (option D) provides no benefit at no cost. The *benefit-to-cost* ratio, or PD return $r = b/c$, is often used to parameterize the game, taking $c = 1$ as monetary unit (see SI note 1 for other parameterizations). Compared to other social dilemmas, the PD is considered the worst-case for the evolution of cooperation^{5–7} (SI note 2).

The standard way to test whether a cooperative strategy (strategy C) has any chance to evolve, is to confront it with the benchmark strategy 'unconditional defection' (strategy D; agents who always defect) in playing the PD under one or a few evolutionary processes. The three different issues to be discussed are: the *invasion* of the strategy, i.e., the spreading of cooperators (C-agents, or C's) in a population dominated by defectors (D-agents, or D's); its *persistence*, i.e., the long-term stabilization, fluctuating or not, of C's; and its *fixation*, i.e., the convergence to the state all-C. For example, it is well known that when the PD is played in large and well-mixed—*unstructured*—populations, there is no hope for the strategy 'unconditional cooperation' (agents who always cooperate) against

Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, I-20133, Milano, Italy. Correspondence and requests for materials should be addressed to F.D. (email: fabio.dercole@polimi.it)

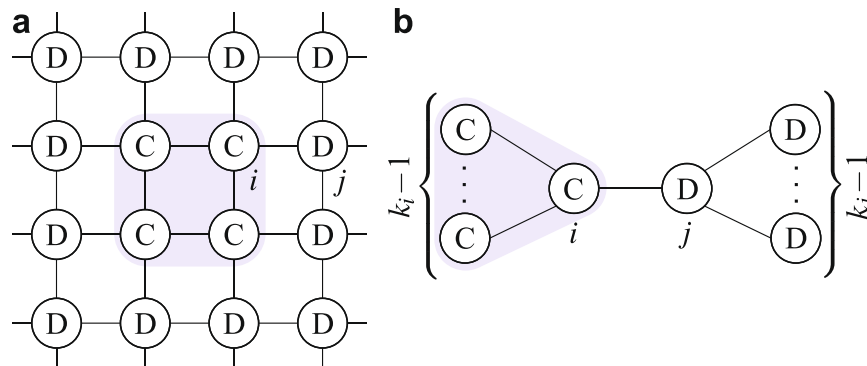


Figure 1. Network reciprocity in regular and non-regular networks. **(a)** A cluster of 4 unconditional C's surrounded by unconditional D's in a square lattice with periodic boundary conditions. The payoff (per game round, obtained by each agent by playing a PD with all neighbors and summing up outcomes) of C's is $\pi_i = 2r - 4$; that of the D's at the boundary of the cluster is $\pi_j = r$. C's do better than D's ($\pi_i > \pi_j$) if $r > 4$, so that, under imitation update, C's remain such and boundary D's change to C (e.g., j copying i), reducing however their payoff to $r - 4$. **(b)** A cluster of k_i C's protected from k_j D's ($k_i, k_j > 2$; k_i and k_j are the degrees of nodes i and j). The payoff of the C-agent i is $\pi_i = r(k_i - 1) - k_i$; that of the D-agent j is $\pi_j = r$; $\pi_i > \pi_j$ if $r > 1 + 2/(k_i - 2)$, so that, under imitation update, i remains C and j drastically reduces her payoff when copying i .

the benchmark D-strategy. Defecting gives the largest payoff regardless of what the others are doing; thus, without any specific incentive to cooperation, C's cannot invade under any reasonable evolutionary process and disappear if initially present in the population.

Traditional incentivizing mechanisms⁹ either make cooperation conditional—such as reciprocal altruism^{1,2} (also known as *direct reciprocity*), the establishment of reputations¹⁰ (also known as *indirect reciprocity*), mechanisms of kin¹¹ or group¹² selection (or other forms of assortment^{13,14}), and the consideration of social^{15,16} and moral¹⁷ preferences—or change the rules of the game, as by introducing optional participation^{18,19} and punishment of antisocial behaviors^{20,21}. All these mechanisms add degrees of strategical complexity, either in terms of players' cognitive abilities and/or information flows, or due to extra options in the underlying game.

Starting with M. A. Nowak and coauthors' influential papers^{22–24}, the fact that interactions in real populations are *structured* according to a network of physical or social contacts has been proposed as the simplest mechanism—requiring no strategical complexity—to explain cooperation. It has been named *network reciprocity* because theoretical models show that if the network is far from all-to-all (technically *sparse*, i.e. with average degree $\langle k \rangle$ —the average number of neighbors—significantly lower than the number N of nodes), it allows C-agents to self-assort into clusters within which they cooperate. That is, the network induces mutual cooperation between C's, even with no (direct) reciprocity built-in in their strategy (Fig. 1). More specifically, the original model of network reciprocity works as follows²⁴. In large regular networks (where each node has k neighbors), driven by an imitative evolutionary process (agents imitate better performing neighbors, see SI note 3), and in the limit of weak selection (the game payoff marginally impacts the individual performance, see SI note 4), unconditional C's can invade unconditional D's and fixate under a simple condition: the PD return r must exceed the connectivity k . This condition has been generalized to non-regular networks^{25,26}, essentially requiring r to exceed the average degree $\langle k \rangle$.

Network reciprocity has been confirmed in several other models^{27–43} (with a few exceptions, see, e.g., ref.⁴⁴). In contrast, the experimental evidence among humans is controversial (see ref.³ for a review). Some authors^{45–50} questioned network reciprocity as a mechanism supporting human cooperation, because experiments on several networks (including a few impressively large regular and non-regular sparse networks^{45,47}) gave low levels of cooperation, similar to those observed with all-to-all interaction. However, most of these experiments failed to satisfy the condition $r > \langle k \rangle$, and more recent experiments satisfying the condition showed higher levels of cooperation, in accordance with the theoretical prediction^{51,52}.

There is nonetheless another argument fueling the apparent mismatch between theory and practice. While all models confirming network reciprocity assume an imitative process of strategy update, the in-depth analysis of some of the experiments showed that human subjects did not take decisions by comparing with neighbors' payoffs^{46,50}. Imitation indeed lacks rationality when interaction patterns are limited and possibly heterogeneous: Why should we copy a better performing neighbor whose neighborhood might be considerably different—in size as well as in composition—from ours? Especially in heterogeneous networks (networks encompassing nodes with very low and high degree), imitation may turn counterproductive (see Fig. 1, where agent j reduces her payoff by copying i). Not surprisingly, by changing the rule for strategy update from imitation to 'best response'—the simplest, though myopic, rational rule of doing what is best for oneself in the next game round—network reciprocity no longer works⁵³, as unconditional defection dominates the PD in any network.

In conclusion, the rule $r > \langle k \rangle$ of network reciprocity apparently works, but the underlying theoretical justification seems not. To help resolving the controversy, we design a model that shows network reciprocity, while being more adherent to the common rationale and to the available experiments than previous models based on

Parameter	Description	Reference values	Other values
N	network size	1000	—
r	PD return	[1, 6]	[20, 60] in Fig. S2e
δ	rate of strategy update	0.05	0.025, 0.1 in Fig. S3 (right)
h	predictive horizon	2, 3, 4, 5	—
d	direct reciprocity ($d_{\min} = -(1/\delta - 1)$; $d_{\max} = 1$)	0 (normal)	1/2 (super), -1 (sub) in Fig. S3 (left)
$\delta_d = (1 - d)\delta$	reciprocity-biased update rate	0.05	0.025, 0.1 in Fig. S3 (left)
$\langle a \rangle$	average length of abstention periods	4.56	6.90, 2.91 in Fig. S3

Table 1. Model parameters (first part) and related quantities (second part).

imitation update. Before describing our model, let us be more precise on what we mean by the fact that an EGT setup, whether theoretical or experimental, shows network reciprocity. We think it is worth to conceptually separate the network from the, possibly induced, reciprocity. The network is the mechanism limiting agents' interactions. Whether this mechanism does enhance mutual cooperation—reciprocity—or not, with respect to the case of all-to-all interaction, evidently depends on the setup details. For example, the answer is 'yes' for the PD played, under $r > \langle k \rangle$, by unconditional C's and D's in models based on imitation; 'no' if the update rule is changed to best response. In other words, we say that the EGT setup shows network reciprocity if a condition qualitatively similar to $r > \langle k \rangle$ holds true, i.e., if cooperation invades, persists, or even fixates in the population, provided the game return is sufficiently large relative to a measure of network connectivity. In this line of thinking, network reciprocity is evidently an outcome of the EGT setup, not a mechanism built-in in the setup itself.

We do not aim at fitting any of the available experiments. Actually, what we take from them is the pairwise PD interaction, mostly used, and three general emerging facts. The first is that humans are conditioned by their current 'mood' to be cooperative or defective; the second is that, while in the C-mood, they tend to reciprocate cooperation; the third is that they defect rather unconditionally while in the D-mood. This evidence justifies our assumption of two strategies, a C-strategy implementing a form of direct reciprocity, and the benchmark D-strategy. We implement direct reciprocity by giving C-agents a temporary exit option^{18,54} toward neighboring defectors, so to reduce exploitation risk. As in the famous tit-for-tat strategy, traditionally used to model direct reciprocity², C's cooperate with neighbors known to cooperate. However, they do not retaliate for defection; rather, they abstain from playing for a few rounds; eventually, they forgive defection and go back playing to seek cooperation, i.e., they poll previous exploiters for a change of mood. Abstention has no cost and gives no benefit to both opponents of the interaction, and we assume it can be distinguished from defection. This is because we imagine interactions taking place in a physical or virtual space in which one notes whether the opponent shows up or not, though identities might not be disclosed. This means that, by abstaining, an agent communicates her C-mood to the opponent, provided the latter shows up. At the same time, abstention provides no information on the opponent's mood.

The rule for strategy update is where we embed rationality in our model. Following the standard EGT tenet of selfishness, we extend the best response rule to the prediction of future incomes over a multi-step horizon of $h \geq 2$ future interactions. The prediction is based on our model, that sets the rules of the society and that is assumed to be common knowledge to all agents. The only further available information directly comes from the pairwise interactions. No access is given to neighbors' payoffs and connectivity. Each agent revises her current strategy from time to time, independently of others, in accordance with a rate δ assumed uniform across the population. Technically, δ is the probability of strategy update after each game round; $1/\delta$ measures the agents' 'inertia' to change (the average number of rounds between two consecutive updates by the same agent). When revising the strategy, an agent computes the cumulated income expected from using the C and D strategies over the predictive horizon, and selects the best one up to the next update. Because of the limited information, the prediction cannot account for changes in the neighbors' strategies. This limits the horizon to a few rounds under a relatively slow strategy update (small δ); because of the short horizon, no discount of future incomes is adopted.

Based on the update rate δ , we set the measure for direct reciprocity. We say that 'normally' reciprocating C-agents are those who, at each round of an abstention period, decide whether to go back playing in accordance with the expectation that the neighboring defector has revised her strategy. 'Super/sub'-reciprocating C's abstain for longer/shorter periods, $\langle a \rangle$ on average. They do so by intentionally under/over-biasing their neighbors' update rate to $\delta_d = (1 - d)\delta$, where -100% is the negative/positive bias. Parameter d modulates the strength of direct reciprocity, from no reciprocity at $d = d_{\min} = -(1/\delta - 1)$, yielding $\delta_d = 1$ and the unconditional C strategy, to extreme reciprocity at $d = d_{\max} = 1$, yielding $\delta_d = 0$ and C's who definitely break connections with defecting neighbors; $d = 0$ is the unbiased case of normal reciprocity. Correspondingly, we say that direct reciprocity is super/sub-normal if d is positive/negative.

We close this introduction by anticipating that this particular form of direct reciprocity is not crucial for our results. We use it because we consider abstention, whether possible, more connatural to the cooperative mood than retaliation. This issue is addressed in the discussion, where we further comment on the novelty and scope of our model. Overall, the model-predictive rule for strategy update is the new key element, that we use to describe a rational decision making. Further details on the model's implementation are given in the Methods. The model parameters are summarized in Table 1.

Results

Our main result is that cooperation fixates in our model starting from any cluster of at least two C-agents, provided that the PD return r is larger than a fixation threshold R_{fix} that increases with the maximal degree k_{max} in the network. The threshold R_{fix} is derived analytically and is very conservative. It represents the worst-case with respect to the position of the initial C's in the network. In practice, the fixation threshold is more related to the network's average degree $\langle k \rangle$. This is shown numerically in Fig. 2, in which the actual threshold r_{fix} is identified by averaging over many simulations, and shown to increase by doubling the average degree from 4 to 8 in several network structures (planar lattices; single-scale: networks with narrow degree distribution; scale-free: broad degree distribution; see Sect. Numerical results for further details). We hence rediscover the rule of network reciprocity over social ties, explained however by direct reciprocity and model-predictive rationality.

Secondary results of our model, proved analytically and quantitatively confirmed numerically, are listed below.

1. Direct reciprocity ($d > d_{\text{min}}$) is essential for the evolution of cooperation. Moreover, the stronger is reciprocity (larger $d \in (d_{\text{min}}, d_{\text{max}})$), the better is for cooperation, in the sense of a reduced fixation threshold.
2. The multi-step prediction ($h \geq 2$) is essential as well, and extending the horizon (larger h) helps cooperation.
3. Cooperation can invade, even starting from a single C-agent. With a single C, the fixation granted by the threshold R_{fix} occurs with a probability that grows from 1/2 to 1 with the degree of the initial C. This is particularly interesting, as models based on imitation update require a significant initial fraction of C's to stabilize cooperation (the issue is further addressed in the Discussion).
4. The degree heterogeneity of the network does not help cooperation, unless the initial C's occupy the network's hubs (Fig. 2: compare the same networks under the two types of placement—random and degree-rank—of the initial C's). Both these effects turned out to be rather limited, in agreement with experimental observations^{47,48}, while in sharp contrast with the theoretical predictions based on imitation update.

Analytical results. In this section we present the properties of our model's dynamics that we proved analytically (some of the proofs are reported in the Supplementary Information). We preliminary consider the case of an infinite predictive horizon. Although lacking sense in our model (because predictions do not account for changes in the neighbors' strategies, see Introduction), it allows to derive a simple condition for the fixation of cooperation. We prove (in SI Sects S1–S6) that when a C-agent with degree k and k_C C-neighbors (known from past interactions) revises her strategy according to an infinite horizon, she remains C (because she expects a loss in changing to D) if and only if

$$r > 1 + \frac{k}{k_C} \frac{P_{\text{CD}}^{\infty}}{1 - P_{\text{CD}}^{\infty}}, \quad P_{\text{CD}}^{\infty} = \frac{1}{2} \frac{\sqrt{4\delta_d - 3\delta_d^2} - \delta_d}{1 - \delta_d} \simeq \sqrt{\delta_d} \text{ for small } \delta_d, \quad \delta_d = (1 - d)\delta \in (0, 1). \quad (1)$$

P_{CD}^{∞} is the probability (computed in SI Sect. S2) that a C-agent who remains C forever will get exploited in a far-future interaction by a D-neighbor who remains D forever. It only depends on the reciprocity-biased update rate δ_d and increases from zero to one with δ_d (SI Fig. S8).

Similarly, the D-to-C strategy change (expected gain in changing to C) occurs under the same condition (1). The harshest condition in (1), i.e., the one for $k = k_{\text{max}}$ and $k_C = 1$, hence gives the threshold R_{fix} of our main result, in the case of an infinite horizon:

$$R_{\text{fix}}^{\infty} = 1 + k_{\text{max}} \frac{P_{\text{CD}}^{\infty}}{1 - P_{\text{CD}}^{\infty}}. \quad (2)$$

With a finite horizon of h future interactions, the conditions governing strategy update are more complex. In SI Sects S4 and S5, we compute the gains $\Delta\pi_C^h$ and $\Delta\pi_D^h$ (positive for a gain; negative for a loss) respectively predicted by a C and a D for changing strategy. The resulting expressions are reported in the Methods. Given a C and a D with identical neighborhoods (including $k_C \geq 1$ C's), the r -threshold for a C to remain C and that for a D to change to C are different. Typically, the former is lower, because the C-neighbors of a C more likely cooperate in the near future than the C-neighbors of a D.

Not surprisingly, for $h = 1$ (best response update) we have $\Delta\pi_C^1 > 0$ and $\Delta\pi_D^1 < 0$ independently of the PD return r and of the neighborhood's composition, so that defecting assures the highest payoff. However, for any $h \geq 2$ and provided the agent has a C-neighbor ($k_C \geq 1$), the expected gains $\Delta\pi_C^h$ and $\Delta\pi_D^h$ depend linearly on r , respectively decreasing and increasing with positive and negative values at $r = 1$. This shows the existence of the threshold R_{fix} , an upper bound of which is provided in SI Sect. S7 by considering the network's node with maximal degree.

Finally, in SI Sect. S8, we further analyze the role of the predictive horizon. Under a condition on r more restrictive than (1), the expected gains $\Delta\pi_C^h$ and $\Delta\pi_D^h$ depend monotonically on h , respectively decreasing and increasing to the negative and positive infinite-horizon limits. For intermediate r , $\Delta\pi_C^h$ (resp. $\Delta\pi_D^h$) first increases (decreases) with h up to a positive (negative) extremum, then decreases (increases) to the negative (positive) infinite-horizon limit.

Despite the system's complexity, the above results have the following consequences, that include our secondary results.

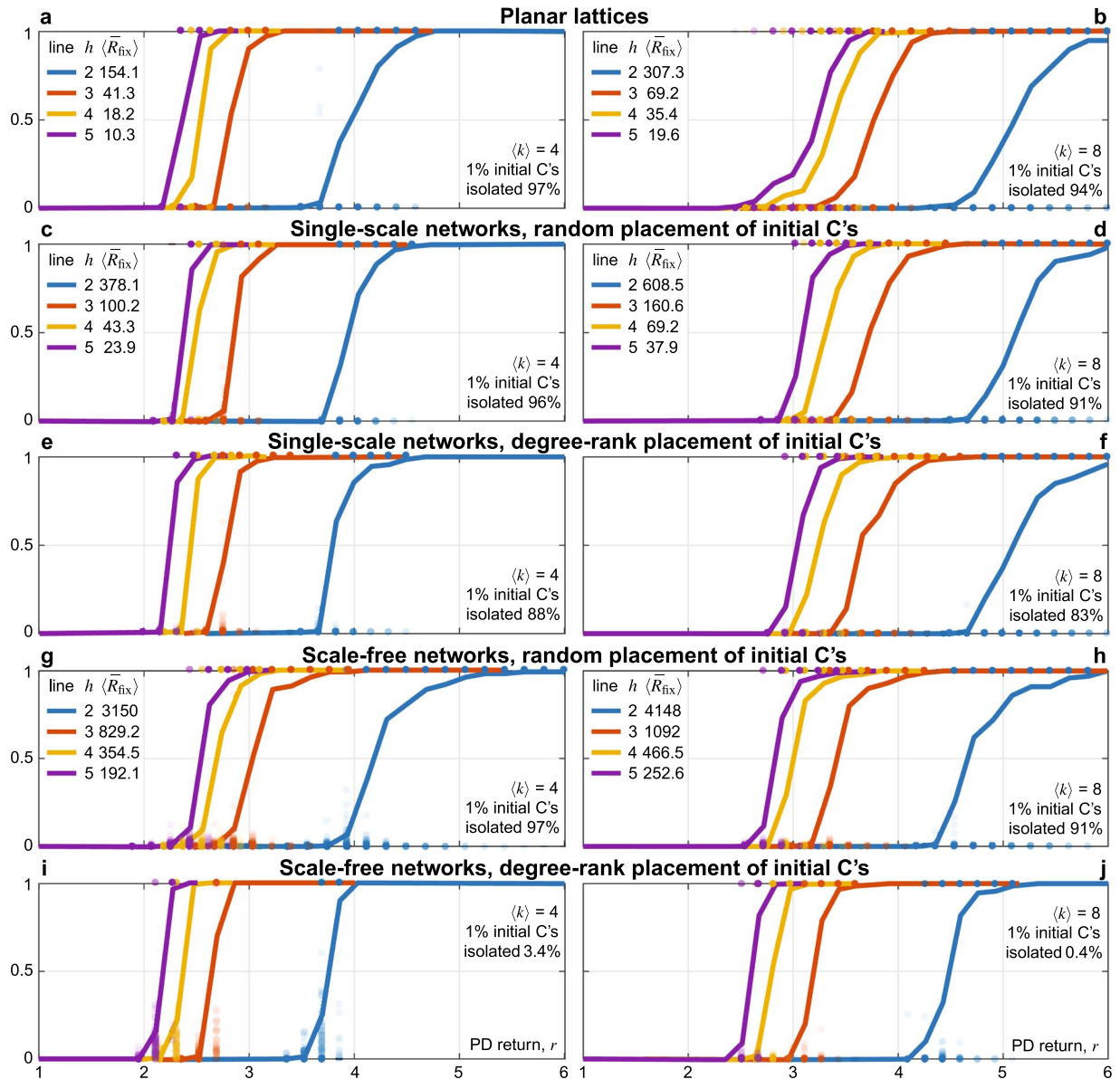


Figure 2. Invasion, persistence, and fixation of cooperation under direct reciprocity and model-predictive rationality. Panels show the level of cooperation reached in 10^4 game rounds starting from 1% initial C's on different network structures (left: average degree $\langle k \rangle = 4$; right: $\langle k \rangle = 8$) as a function of the PD return r . Solid lines show the average fraction over 100 random initializations (random placement of the initial C's in planar lattices; network generation and random placement of the initial C's for random—single scale and scale free—networks; the average % of isolated initial C's is reported). Dots show the outcome of single simulations (only for $r_{\text{inv}} < r < r_{\text{fix}}$, i.e., only if some of the outcomes lie in the open interval $(0, 1)$); transparency is used to show dots accumulation. Colors code the predictive horizon h , from 2 to 5, and the corresponding upper bound \bar{R}_{fix} to the threshold R_{fix} is reported. Model parameters: reference values in Table 1. See Sects S11 and S12 for further details on networks and numerical simulations.

- (i) The state all-C is an equilibrium of our model (or better an absorbing state of the Markovian dynamics) under a requirement on r that is much milder than the fixation threshold (see condition (1) with $k_C = k$).
- (ii) The state all-D is always an equilibrium (see condition (1) with $k_C = 0$), though a coordinated switch to C by a small cluster of players (not considered in our model) can give a payoff gain.
- (iii) Isolated C's ($k_C = 0$) change to D as soon as they revise their strategy. This does not mean that cooperation cannot start from a single C-agent, since D-neighbors could change to C before the isolated C changes to D.
- (iv) Indeed, a D player connected to an isolated C can change to C, provided r is sufficiently large (see condition (1) with $k_C = 1$).
- (v) Secondary result 1: direct reciprocity ($d > d_{\text{min}}$) is necessary for the evolution of cooperation. With no reciprocity ($d = d_{\text{min}}$), the C-strategy is unconditional ($\delta_d = 1$ and $P_{\text{CD}}^\infty = 1$ in (1)) and D is the well-

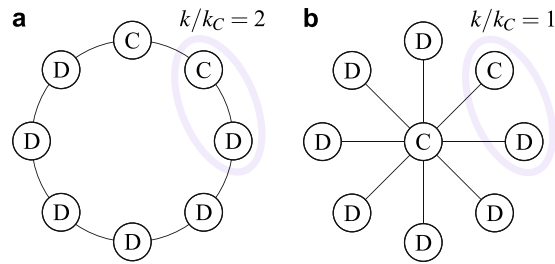


Figure 3. Direct reciprocity and model-predictive rationality in regular and heterogeneous networks. Consider (a) the ring ($k = 2$ for all nodes) and (b) the star network ($\langle k \rangle \simeq 2$ for large N) in the infinite-horizon limit. The ring- R_{fix}^{∞} is given by (1) with $k/k_C = 2$ and a single D drives the population to all-D if $r < R_{\text{fix}}^{\infty}$. The star- R_{fix}^{∞} is much higher, because the ratio k/k_C peaks at $N - 1$ for the central node with only one C-neighbor. However, if the central node is a C, the D-leaves will change to C under the weakest requirement on r ($k/k_C = 1$ in (1)) and there are high chances that this occurs before the central node revises her strategy. (The probability that a D-leaf revises before the central node is given by the formula at point (ix), in which k is replaced by the number of D-leaves.) If the number k_C of C-leaves raises to satisfy condition (1), the population then evolves to all-C, and this occurs with probability higher than $1/2$ starting with no C-leaves and r equal to the ring- R_{fix}^{∞} (the probability goes as $1/2 + 1/(2N)$ for small δ , see SI Sect. S9). That is, on average, the isolated central C drives the star to all-C under a milder condition on r w.r.t. the ring with some initial D's.

known best option for any r . Moreover, as P_{CD}^{∞} gets smaller together with δ_d , a stronger reciprocity (larger $d \in (d_{\text{min}}, d_{\text{max}})$) makes the fixation threshold (2) milder.

- (vi) Secondary result 2 (first part): the multi-step predictive horizon ($h \geq 2$) is necessary for cooperation, because of the above discussion on $\Delta\pi_C^1$ and $\Delta\pi_D^1$.
- (vii) Secondary result 2 (second part): increasing the horizon h reduces the threshold R_{fix} . Indeed, if $r = R_{\text{fix}}$ for a given h , the node with maximal degree gives either $\Delta\pi_C^h = 0$ and $\Delta\pi_D^h > 0$ or $\Delta\pi_C^h < 0$ and $\Delta\pi_D^h = 0$, depending on the composition of its neighborhood. By the behavior of $\Delta\pi_C^h$ and $\Delta\pi_D^h$ described above, the former decreases while the latter increases by adding one prediction step. With $h + 1$ steps, then, all C's remain C and all D's changes to C when revising strategy, so that r is above threshold.
- (viii) The inertia to change also helps cooperation, in the sense that, for a given direct reciprocity d , a lower update rate δ gives a lower biased rate δ_d (longer abstention periods, on average), and hence a smaller P_{CD}^{∞} and a milder fixation threshold (2).
- (ix) Secondary result 3: our main result probabilistically holds also starting from a single C in the network, provided her degree k is not too small. With $r > R_{\text{fix}}$, the probability that a D-neighbor changes to C before the initial C changes to D goes as $1 - 1/(k + 1)$ for small δ (shown in SI Sect. S9).
- (x) Secondary result 4 (first part): according to condition (1) and the fixation threshold (2), and given the network size N , cooperation is favored in homogeneous (regular or single-scale) sparse networks, because of the low maximal degree, compared to dense or heterogeneous networks. However, keeping the same average degree, heterogeneous networks have more low-degree nodes compared to homogeneous ones, so that, while isolated C's might have higher chances to change to D (see point (ix)), their D-neighbors change to C, with similar probability, under milder returns (condition (1)). Overall, network heterogeneity should facilitate the invasion of cooperation and make fixation more demanding, provided the initial C's are placed at random.
- (xi) Secondary result 4 (second part): the comment at point (ix) and two simple examples in Fig. 3 suggest that network heterogeneity helps cooperation if C's initially occupy the network's hubs. Essentially, low-degree D's connected to a C change to C under a mild requirement on r (low ratio k/k_C in (1)). However, to have many low-degree D's connected to few initial C's, we need high-connected C's. Hence, especially starting at low levels of cooperation, degree heterogeneity and the placement of the initial C's in the network's hubs together reduce the requirement on r to fixate cooperation. Compared to imitation update (in which C-hubs need a significant fraction of C-neighbors to persist), our predictive strategy update allows the formation of C-clusters even starting from isolated C-hubs, who pay (or better invest in) the initial cost of exploitation. If however most of the hubs are D's, network heterogeneity turns harmful to cooperation (see point (x)).

Numerical results. To quantify the analytical results, we have run many numerical simulations on several networks of $N = 1000$ nodes: ring and planar lattices, single-scale (Watts-Strogatz model with full rewiring) and scale-free (Barabási-Albert model) random networks, and the complete (all-to-all) network. The results for 1% initial fraction of (normally reciprocating, $d = 0$) C's on planar 4- and 8-neighbor lattices and on random networks with average degree $\langle k \rangle = 4$ and 8 are reported in Fig. 2. See SI Fig. S1 for 50% initial C's and S2 for a degree-4 ring lattice (a ring of nodes each connected to the 4 nearest nodes in the loop) and the complete network; see also SI Fig. S3 for the case of super/sub-normal reciprocity (d positive/negative) and for different values of the update rate δ . For random networks, we have separately simulated the random placement of the initial C's and the placement according to degree-ranking.

The simulations confirm that cooperation is able to invade, persist, and fixate in any network, provided the PD return r is large enough. For any combination of network structure, initialization, and model parameters, the simulations starting from 1% initial C's identify two thresholds on r : a lower threshold r_{inv} above which, on average, cooperation invades and persists; an upper threshold r_{fix} above which cooperation always invades and fixates. For r between r_{inv} and r_{fix} , the average asymptotic fraction of C's (solid lines) is not representative of the level of cooperation one should expect in a single simulation (dots), as cooperation disappears/fixates in most of the cases (see dots at fractions 0 and 1) and typically settles at low C-levels (below 20%) in the rest of the cases. A deeper analysis of the simulations of Fig. 2 not ended in all-C or all-D (the dots in the open interval (0, 1)) indeed reveals that most of dots above 0.2 denote simulations that converge to all-C on a longer timescale, whereas dots below 0.2 typically represents simulations ending in a nontrivial stalemate—an equilibrium different from all-C and all-D—or showing long-term fluctuations (see SI Sect. S10 for examples of nontrivial stalemates and fluctuations in the simple network of Fig. 1b; Sects S11 and S12 for further details on networks and numerical simulations).

Note that the threshold r_{fix} is much smaller than the theoretical R_{fix} of our main result (see the average $\langle \bar{R}_{\text{fix}} \rangle$ over the simulations of a given type in Fig. 2, where \bar{R}_{fix} is the upper bound to R_{fix} derived in SI Sect. S7). This is due to the stochastic effect described at point (ix). Essentially, even if some of the C's initially need a higher r to remain C, by the time they revise their strategy the r -gap could have vanished because of D-to-C changes in the neighborhood. This overcompensates the opposite effect due to the fact that R_{fix} is computed starting from a cluster of two C's, whereas initial C's are most often isolated (except for scale-free networks with degree-rank-C-placement, because hubs are likely connected among themselves; see the average % for each panel in Fig. 2). Starting with random pairs of connected initial C's indeed results in lower r_{inv} and r_{fix} (shown in SI Fig. S4).

As expected from the arguments at points (vii) and (x), the thresholds r_{inv} and r_{fix} decrease if the horizon h is extended and increase with the network's average degree, given all other details (Fig. 2: compare the different colors within each panel and left vs right panels). Both effects are weakened, as expected, starting from higher initial C-levels (shown in SI Fig. S1).

Degree heterogeneity also works as theoretically predicted. It works against the fixation of cooperation under random placement (Fig. 2: compare single-scale and scale-free networks and note the significantly higher r_{fix} in the latter; r_{inv} is slightly lower, as predicted at point (x)). The effects are again weakened starting from higher initial C-levels (SI Fig. S1). Placing the initial C's in the network's hubs does favor cooperation, both in terms of invasion and fixation (Fig. 2: compare single-scale and scale-free networks under degree-rank-C-placement and note the lower r_{inv} and r_{fix} in scale-free networks; also note that the type of placement is irrelevant for single-scale networks). The effect is however moderate.

Finally, we have checked the robustness of the above results with respect to a different type of scale-free network (with tunable transitivity, Holme-Kim model, see SI Fig. S5), and by limiting the agents' computational skills and rationality (see SI Figs S6 and S7, respectively).

Discussion

We have designed and analyzed an EGT model to validate two specific hypotheses: (1) that reciprocal cooperation can evolve on social networks, against unconditional defection, under a process of strategy update that rationally pursues the individual interest; and (2) that this EGT setup shows network reciprocity. Our motivation originates from a mismatch between theoretical predictions and experiments with human subjects. Network reciprocity—the beneficial effect on cooperation possibly emerging when agents' interactions are limited by a sparse network of contact—has been theoretically investigated with models based on imitative processes of strategy update, with the result that cooperation has better chances to fixate, the fewer are, on average, the connections between the agents. This 'rule' of network reciprocity, precisely the condition $r > \langle k \rangle$, where r is the economic return of the game interaction and $\langle k \rangle$ is the network's average degree, is consistent with all available experiments^{45–47,51,52,55–58}, in which however the subjects' behavior proved incompatible with the assumed imitation process. Indeed, when the interaction is far from all-to-all, imitating a better performing neighbor gives no guarantee to increase our own payoff, and models implementing the simplest rational rule to do what is best for us in the next interaction—the best response update—show no network reciprocity⁵³. Our results reconcile the mismatch, showing network reciprocity in an EGT setup that is more adherent to the common rationale and to the behavioral traits emerging from experiments.

Our model does not specifically describe any of the available experimental setups. The link to the experiments is in the choice of the two strategies we confront: a C-strategy implementing a form of direct reciprocity and the benchmark D-strategy of unconditional defection. The analysis of some of the experiments (in which, as a rule of the experiment, subjects take one decision per game round to cooperate or defect with all neighbors) indeed revealed three general facts^{45,47,48,50}: the first is that subjects' behavior is typically biased by previous decisions (cooperation being more/less likely in subjects who cooperated/defected in the previous round), so that a cooperative or defective 'mood' can be identified. The second is that, while in the C-mood, subjects reciprocate cooperation (cooperation being more/less likely the higher/lower was the number of cooperating neighbors in the previous round); third, subjects in the D-mood defect rather unconditionally.

The main novelty of our model is the predictive rule for strategy update, that extends the best response rule to an horizon of future interactions. It is a modeling assumption intentionally not grounded on experimental evidence; it is our way to introduce rationality in the model. Identifying the update rule from experiments is in any case a difficult task^{46,48,50,57} and, somehow, an ill-posed problem. The neat distinction between strategies and strategy update is actually pertinent to models only; e.g., the observed moody behavior has been also described as a unique strategy conditioned by the agent last decision⁴⁸. Moreover, different subjects can use (or learn) different rules and the result is likely to depend on the experimental setting. Although the average rates at which subjects

changed between the C and D moods have been estimated from data^{45–48,50}, and turned out independent on the network structure⁴⁶, this does not unveil the decision-making process.

Note that the novelty is not strictly in the use of a non-imitative, also called *innovative*, rule for strategy update, as the best response is itself innovative. Moreover, there has recently been a growing interest for innovative rules^{59,60} and for the coexistence⁶¹ and competition⁶² between imitative and innovative rules. So far, however, no innovative update showed any significant network effect on cooperation^{53,59}, at the point that network reciprocity is often considered a feature of imitative dynamics only⁶¹. The novelty is hence to show network reciprocity under an innovative evolutionary process.

The way in which we implement direct reciprocity—allowing C's to selectively abstain from playing the PD with D-neighbors—also brings some novelty. Optional participation is known to relax social dilemmas when a baseline payoff is granted to loners^{19,63,64}. However, abstention has often been considered an independent strategy^{63–65}, or a probabilistic option for both C's and D's⁶⁶, rather than an option for cooperators to lower exploitation risk (as proposed in ref.¹⁹ in a non-evolutionary setting). This is the option that makes abstention a form of direct reciprocity. Differently from previous models, we assign no payoff to such an option. We prefer abstention rather than retaliation—cooperators defecting exploiters—because this is more connatural to the C-mood. Although there is typically no difference in a single round (because both abstention and mutual defection give no payoff), abstaining C's communicate their mood to exploiters, who will take it into account when revising their strategy. This is because we assume that agents who agree to play see whether the opponent agrees or abstain, whereas abstaining agents get no information. Of course our model does not represent situations in which the interaction is blind and agents only see their own payoff.

Direct reciprocity deserves another comment. It requires repeated interactions among the same agents, as well as cognitive abilities to recognize individuals and remember past interactions. Originally^{1,2}, it has been studied in iterated games, i.e., (non-evolutionary) games involving only two players (rather than a population of two types of players) who know the probability $w > 0$ of a next interaction. In EGT, there are two ways in which one can study direct reciprocity: either the single game round consists of an iterated game between each pair of neighbors—an option allowing direct reciprocity in any network, so far investigated in the socio-economic context under imitative strategy update^{32,67}—or the single game is one-shot, but agents repeatedly interact on a static (or slowly changing) sparse network, as we do. Indeed, the unchanging network grants a next round, while the sparsity of the connections makes cognitive tasks affordable. Of course hubs need more resources than leaves, but this is typically built-in in the socio-economic structure of the society.

The other elements of our model are rather standard (see Methods for full detail). We now discuss our results and their scope in the light of previous theoretical and empirical work. First of all, the interplay between direct and network reciprocity has been already addressed^{32,67} with models that, in the socio-economic context, are however based on imitative strategy update. The general result is that direct reciprocity in the C-strategy favors cooperation in synergism with the population structure, the latter being defined by a network³² or through assortative encounters⁶⁷. Our model confirms the role of direct reciprocity in triggering the positive network effect on cooperation, and does it under a rational process of strategy update.

We actually show that direct reciprocity, or some other mechanism supporting cooperation (see the Introduction), is necessary to rationally avoid the dominance of defection in a networked PD, and hence to see any network effect. Focusing on direct reciprocity, we expect similar results by rationally confronting unconditional D with any reciprocating form of conditional cooperation. To test this claim, we have run preliminary simulations on the famous tit-for-tat strategy² (traditionally used to model direct reciprocity, including refs.^{32,67}). Worth of note is that our model highlights a double role of static and sparse networks of contacts. Not only these are the structures that favor cooperation according to our model, but these are, as well, the structures that make our model feasible, because the cognitive tasks required by both direct reciprocity and model-predictions scale with the size of the agent's neighborhood. We did use our model on a large complete network as well, but we did it for benchmarking purposes, to set the case to compare with in looking for network effects.

Should we then expect to see a decline of human cooperation in real-world socio-economic systems that are arguably getting more and more connected? This seems to be the take-home message from network reciprocity, that our model even made more convincing. Models and experiments showing network reciprocity were however designed to isolate the effect of static network of contacts. In real-world systems, other mechanisms can support cooperation despite the high interconnection, such as cultural, economic, and political agreements, or punishment mechanisms; the latter, in particular, has been recently shown to inhibit the network effect in experiments⁵². Moreover, real-world networks are often dynamic. In particular, adaptive networks, in which agents can cut links and establish new ones as a result of the game interaction, proved to be cooperation promoters, with both theoretical support^{65,68,69}, however based on imitation update, and empirical evidence^{70–72} (see ref.⁷³ for a review). Note that allowing C-agents to cut links with exploiters and seek for new cooperators is an exit option^{18,54} similar to the one we use; we have a static topology of connection, but links are temporarily inhibited. Interestingly, adaptive rewiring can be feasible even in medium-large neighborhoods, as, e.g., random rewiring does not involve intense cognitive tasks. It would be then interesting to study adaptive rewiring under model-predictive update, to see whether it supports the evolution of cooperation in dense networks.

Two aspects on which our results significantly differ from those obtained with imitative update concern the invasion potential of cooperation and the role played by the degree heterogeneity of the network. About the first aspect, we have shown (analytically and quantified numerically) the existence of a threshold on the game return r (scaling with the network connectivity) above which cooperation has high chances to fixate starting from a single C-agent, i.e., chances of the order $1 - 1/(k + 1)$ at a low rate of strategy update, where k is the degree of the initial C (see analytical result (ix)). This is different from what is granted by imitative update under the condition $r > \langle k \rangle$, i.e., the fixation probability—the probability that cooperation fixates starting from a single randomly

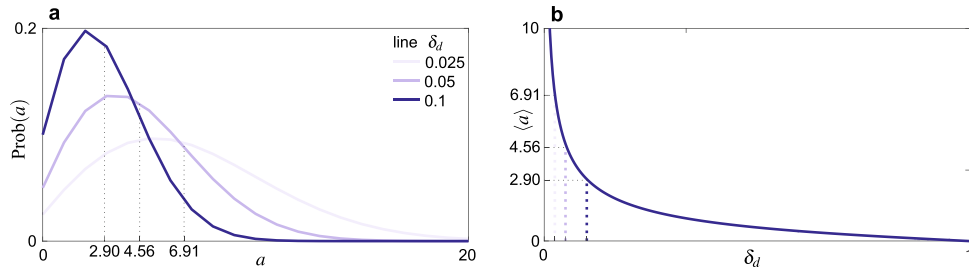


Figure 4. The statistics of abstention periods. **(a)** The distribution Prob(*a*) of Eq. (4). **(b)** The mean $\langle a \rangle$ of the distribution as a function of its parameter δ_d .

The matrix gives the payoffs for the row agent; those for the column agent are given by the transposed matrix. The shading of the first row indicates that those zero payoffs are indistinguishable by an abstaining C-agent, whereas, a D-agent (third row) is able to distinguish between abstaining C and D neighbors.

The abstention mechanism implements a form of direct reciprocity between C-agents. After getting exploited by the neighbor *j*, the C-agent *i* decides to abstain at the next game round with the probability $1 - \delta$ that *j* has not revised strategy. Similarly, after $t - 1$ consecutive abstentions, *i* opts for another one with the probability $(1 - \delta)^t$ (vanishing for increasing t) that *j* has not revised since having exploited *i*. This is the behavior of ‘normally’ reciprocating C-agents (normal reciprocity). ‘Super/sub’-reciprocating C’s abstain for longer/shorter periods, on average (super/sub-normal reciprocity). They formally behave as normally reciprocating ones, but base their decisions on the under/over-biased update rate δ_d defined in (1), where -100% is the negative/positive bias, also assumed uniform across the population. Parameter $d \in (d_{\min} = -(1/\delta - 1), d_{\max} = 1)$ modulates the strength of direct reciprocity: no reciprocity at $d = d_{\min}$; sub-normal reciprocity for $d \in (d_{\min}, 0)$; normal reciprocity at $d = 0$; super-normal reciprocity for $d \in (0, d_{\max})$; extreme reciprocity at $d = d_{\max}$ (the extreme cases $d = d_{\min}$ and $d = d_{\max}$ are considered unadmissible in our model). The resulting distribution of the length *a* of abstention periods, i.e., the probability of *a* consecutive abstentions followed by a cooperation, is

$$\text{Prob}(a) = \begin{cases} \delta_d & \text{if } a = 0, \\ \prod_{t=1}^a (1 - \delta_d)^t (1 - (1 - \delta_d)^{a+1}) = (1 - \delta_d)^{a(a+1)/2} (1 - (1 - \delta_d)^{a+1}) & \text{if } a > 0. \end{cases} \quad (4)$$

It is graphed in Fig. 4 for different values of the reciprocity-biased update rate δ_d , together with the δ_d -dependence of its mean.

We implement our model by endowing each agent *i* with the set of probabilities p_{ij} ’s, $j = 1, \dots, N$, that *i* will agree to play with *j* at the next game round. Initially, $p_{ij} = p_{ji} = 1$ if *i* and *j* are neighbors, while $p_{ij} = p_{ji} = 0$ otherwise; the $N \times N$ matrix $P = [p_{ij}]$ hence defines the network topology. At each round, the (*i*, *j*) PD interaction takes place with probability $p_{ij}p_{ji}$, i.e., only if both players agree to play. When the C-agent *i* gets exploited by neighbor *j*, she sets $p_{ij} = p_{ji} = \delta_d$. If *i* decides not to play with *j* at the next round, p_{ij} is updated to $p_2 = 1 - (1 - \delta_d)^2$; after $t - 1$ consecutive abstentions, the probability to play at the next round is

$$p_t = 1 - (1 - \delta_d)^t, \quad t \geq 1, \quad (5)$$

i.e., the probability (increasing to one with t) that *j* has revised at least once since having exploited *i*. As soon as *i* decides to play, p_{ij} is reset to p_1 if *j* defects again, otherwise it is set to $p_0 = 1$ to reciprocate cooperation. D’s always have $p_{ij} = 1$ toward all neighbors. We therefore implement a static network topology with dynamic weights (the probabilities p_{ij}) associated to the connections.

Strategy update is asynchronous⁷⁵, with rate $\delta \in (0, 1)$ (per game round) uniform across the population, i.e., after each game round (also following the update of the probabilities p_{ij}), each individual independently decides whether to revise her strategy with probability δ . C’s who revise compute the payoffs they expect to collect if remaining C, π_{CC}^h , or behaving as D, π_{CD}^h , over the next h rounds, and change to D if the expected gain $\Delta\pi_C^h = \pi_{CD}^h - \pi_{CC}^h$ is positive. Similarly, D’s who revise compute their expected collected payoff π_{DC}^h and π_{DD}^h and change to C under a positive expected gain $\Delta\pi_D^h = \pi_{DC}^h - \pi_{DD}^h$. Future payoffs are not discounted because of the short predictive horizon h (see Model parameters below). When changing to D, the C-agent *i* sets $p_{ij} = 1$ toward all neighbors. When changing to C, D-agent *i* sets $p_{ij} = p_{ji} = \delta_d$ toward her D-neighbors, to possibly avoid being exploited at the next round. Note that, by construction, either $p_{ij} = 1$, or $p_{ji} = 1$, or both.

The payoff predictions are computed assuming that neighbors behave according to the model society here described. The information available to agent *i* on her *j* neighbor’s state (strategy and probabilities p_{ji} ’s) only comes from the pairwise interactions and is summarized in the sets of probabilities p_{ij} ’s and p_{ji} ’s. Indeed, for agent *i*, remembering the last (*i*, *j*) PD interaction, i.e., when it took place and the outcome, is equivalent to update both sets of probabilities, according to the assumed abstention mechanism. For the revising C-agent *i*, $p_{ij} < 1$ implies that *j* exploited *i* at the last (*i*, *j*) PD interaction; otherwise $p_{ij} = 1$ and *i* and *j* both cooperated at the last round if

$p_{ji} = 1$ or j did not play because exploited by i at the last (i, j) PD interaction if $p_{ji} < 1$. Agent j is considered a D in the first case, even though she might have changed to C since i and j last played the PD; j is correctly considered a C in the second case. For the D-agent i , $p_{ji} < 1$ implies i exploited j at the last (i, j) PD interaction; otherwise $p_{ji} = 1$ and i and j both defected at the last round. In both cases, agent j is correctly identified as C and D, respectively. Thus, only D-agents have full information on their neighbors' state, whereas C-agents can underestimate the number of their C-neighbors.

Payoff predictions are accordingly computed in SI Sects S4 and S5, disregarding, for simplicity, possible changes in the neighbors' strategies. To this end, two probabilities are first computed in SI Sects S2 and S3 for the revising C-agent i , with $p_{ij} = p_{t_{ij}}$ for some integer t_{ij} to play the PD with the neighbor j , with $p_{ji} = p_{t_{ji}}$ for some integer t_{ji} : the probability $P_{CD}^t(t_{ij})$, $t_{ij} \geq 1$, to play with the D-neighbor j at round t of the predictive horizon; and, similarly, the probability $P_{CC}^t(t_{ji})$, $t_{ji} \geq 0$ to play with the C-neighbor j . They can both iteratively be computed with the following recursions:

$$P_{CD}^{t+1}(t_{ij}) = P_{CD}^t(t_{ij})\delta_d + (1 - P_{CD}^t(t_{ij}))((1 - \delta_d)P_{CD}^t(t_{ij}) + \delta_d), \quad (6a)$$

$$P_{CC}^{t+1}(t_{ji}) = P_{CC}^t(t_{ji}) + (1 - P_{CC}^t(t_{ji}))p_{t_{ji}+t}, \quad (6b)$$

initialized at $P_{CD}^1(t_{ij}) = p_{t_{ij}}$ and $P_{CC}^1(t_{ji}) = p_{t_{ji}}$ (P_{CD}^t converges, as $t \rightarrow \infty$, to the infinite-horizon limit P_{CD}^∞ of condition (1), independently of initialization). The resulting formulas for the expected payoff gains $\Delta\pi_C^h$ and $\Delta\pi_D^h$ are

$$\Delta\pi_C^h = -r \sum_{p_{ij}=1} \underbrace{\left(\sum_{t=1}^h P_{CC}^t(t_{ji}) - \sum_{t=1}^h P_{CD}^t(t_{ji}) \right)}_{>0 \text{ for } h \geq 2} + \sum_{p_{ij}=1, t=1}^h P_{CC}^t(t_{ji}) + \sum_{p_{ij} < 1, t=1}^h P_{CD}^t(t_{ji}), \quad (7a)$$

$$\Delta\pi_D^h = r \sum_{p_{ji} < 1} \underbrace{\left(\sum_{t=1}^h P_{CC}^t(t_{ji}) - \sum_{t=1}^h P_{CD}^t(t_{ji}) \right)}_{>0 \text{ for } h \geq 2} - \sum_{p_{ji} < 1, t=1}^h P_{CC}^t(t_{ji}) - \sum_{p_{ji}=1, t=1}^h P_{CD}^t(1). \quad (7b)$$

Model parameters. The model parameters are summarized in Table 1, together with their reference values and the perturbed values used in the numerical analysis. The size N and the matrix $P = [p_{ij}]$ define the network of the agents' connections; parameters r , δ , h , and d set the dynamics of our model; $\langle a \rangle$ and δ_d are derived quantities. Though we do not have a closed formula for $\langle a \rangle$ as a function of δ_d , there is a one-to-one correspondence (graphically expressed by Fig. 4b) between the two quantities. Either d , δ_d , or $\langle a \rangle$ can therefore be used as a model parameter, the other two being accordingly derived for any given δ .

We consider a slow strategy update (small δ), compared to the discrete time of game rounds. This allows to disregard the neighbors' updates in computing the payoff predictions over relatively short predictive horizons. In our simulations we have limited the product δh —upper bounding the neighborhood fraction possibly subject to change within the horizon—to 0.3 (e.g., $\delta = 0.05$ and $h \leq 5$ in our reference setting of Fig. 2).

We consider significant cases of super- and sub-normal reciprocity: $d = 1/2$ and $d = -1$ corresponding to a -50% and $+100\%$ bias of the update rate used in the statistics of the abstentions. The corresponding average lengths $\langle a \rangle$ of abstention periods is reported in Table 1.

Networks structures and numerical simulations. See Numerical analysis. See SI Sects S11 and S12 for further details.

References

- Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–37 (1971).
- Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
- Perc, M. *et al.* Statistical physics of human cooperation. *Physics Reports* **687**, 1–51 (2017).
- Ocampo-Martinez, C. & Quijano, N. Game-theoretical methods in control of engineering systems: An introduction to the special issue. *IEEE Control Syst.* **37**, 30–32 (2017).
- Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics*. (Cambridge University Press, Cambridge, UK, 1998).
- Nowak, M. A. *Evolutionary Dynamics* (Harvard University Press, 2006).
- Szabó, G. & Fath, G. Evolutionary games on graphs. *Phys. Rep.* **446**, 97–216 (2007).
- Perc, M., Gómez-Gardeñes, J., Szolnoki, A., Floría, L. M. & Moreno, Y. Evolutionary dynamics of group interactions on structured populations: a review. *Journal of the Royal Society Interface* **10**, 20120997 (2013).
- Nowak, M. A. Five rules for the evolution of cooperation. *Science* **314**, 1560–1563 (2006).
- Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577 (1998).
- Hamilton, W. D. The genetical evolution of social behaviour. *J. Theor. Biol.* **7**, 1–16 (1964).
- Wilson, D. S. A general theory of group selection. *Proc. Natl. Acad. Sci. USA* **72**, 143–146 (1975).
- Jansen, V. A. & van Baalen, M. Altruism through beard chromodynamics. *Nature* **440**, 663–666 (2006).
- Fletcher, J. A. & Doebeli, M. A simple and general explanation for the evolution of altruism. *P. Roy. Soc. Lond. B* **276**, 13–19 (2009).
- Bolton, G. E. & Ockenfels, A. A theory of equity, reciprocity and competition. *American Economic Review* **90**, 166–193 (2000).
- Capraro, V. A model of human cooperation in social dilemmas. *PLoS One* **8**, e72427 (2013).
- Capraro, V. & Rand, D. G. Do the right thing: Experimental evidence that preferences for moral behavior, rather than equity or efficiency per se, drive human prosociality. *Judgment and Decision Making* **13**, 99–111 (2018).

18. Vanberg, V. J. & Congleton, R. D. Rationality, morality, and exit. *Am. Polit. Sci. Rev.* **86**, 418–431 (1992).
19. Batali, J. & Kitcher, P. Evolution of altruism in optional and compulsory games. *Journal of Theoretical Biology* **175**, 161–171 (1995).
20. Boyd, R., Gintis, H., Bowles, S. & Richerson, P. J. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* **100**, 3531–3535 (2003).
21. Dercote, F., De Carli, M., Della Rossa, F. & Papadopoulos, A. V. Overpunishing is not necessary to fix cooperation in voluntary public goods games. *J. Theor. Biol.* **326**, 70–81 (2013).
22. Nowak, M. A. & May, R. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).
23. Lieberman, E., Hauert, C. & Nowak, M. A. Evolutionary dynamics on graphs. *Nature* **433**, 312–316 (2005).
24. Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505 (2006).
25. Konno, T. A condition for cooperation in a game on complex networks. *J. Theor. Biol.* **269**, 224–233 (2011).
26. Allen, B. *et al.* Evolutionary dynamics on any population structure. *Nature* **544**, 227–230 (2017).
27. Santos, F. C. & Pacheco, J. M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104 (2005).
28. Santos, F. C. & Pacheco, J. M. A new route to the evolution of cooperation. *J. Evolution. Biol.* **19**, 726–733 (2006).
29. Santos, F. C., Pacheco, J. M. & Tom, L. Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc. Natl. Acad. Sci. USA* **103**, 3490–3494 (2006).
30. Santos, F. C., Rodrigues, J. F. & Pacheco, J. M. Graph topology plays a determinant role in the evolution of cooperation. *Proc. R. Soc. B* **273**, 51–55 (2006).
31. Gómez-Gardeñes, J., Campillo, M., Floría, L. M. & Moreno, Y. Dynamical organization of cooperation in complex topologies. *Phys. Rev. Lett.* **98**, 108103 (2007).
32. Ohtsuki, H. & Nowak, M. A. Direct reciprocity on graphs. *J. Theor. Biol.* **247**, 462–470 (2007).
33. Poncela, J. *et al.* Robustness of cooperation in the evolutionary prisoner's dilemma on complex networks. *New J. Phys.* **9**, 184 (2007).
34. Assenza, S., Gómez-Gardeñes, J. & Latora, V. Enhancement of cooperation in highly clustered scale-free networks. *Phys. Rev. E* **78**, 017101 (2008).
35. Gómez-Gardeñes, J. J. P., Floría, L. M. & Moreno, Y. Natural selection of cooperation and degree hierarchy in heterogeneous populations. *J. Theor. Biol.* **253**, 296–301 (2008).
36. Pusch, A., Weber, S. & Porto, M. Impact of topology on the dynamical organization of cooperation in the prisoner's dilemma game. *Phys. Rev. E* **77**, 036120 (2008).
37. Santos, F. C., Santos, M. D. & Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454**, 213–216 (2008).
38. Devlin, S. & Treloar, T. Evolution of cooperation through the heterogeneity of random networks. *Phys. Rev. E* **79**, 016107 (2009).
39. Devlin, S. & Treloar, T. Cooperation in an evolutionary prisoner's dilemma on networks with degree-degree correlations. *Phys. Rev. E* **80**, 026105 (2009).
40. Floría, L., Gracia-Lázaro, C., Gómez-Gardeñes, J. & Moreno, Y. Social network reciprocity as a phase transition in evolutionary cooperation. *Phys. Rev. E* **79**, 026106 (2009).
41. Perc, M. Evolution of cooperation on scale-free networks subject to error and attack. *New J. Phys.* **11**, 033027 (2009).
42. Cardillo, A., Gómez-Gardeñes, J., Vilone, D. & Sánchez, A. Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. *New J. Phys.* **12**, 103034 (2010).
43. Débarre, F., Hauert, C. & Doebeli, M. Social evolution in structured populations. *Nat. Commun.* **5**, 3409 (2014).
44. Hauert, C. & Doebeli, M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646 (2004).
45. Grujić, J., Fosco, C., Araujo, L., Cuesta, J. A. & Sánchez, A. Social experiments in the mesoscale: Humans playing a spatial prisoner's dilemma. *PLoS One* **5**, e13749 (2010).
46. Grujić, J., Röhl, T., Semmann, D., Milinski, M. & Arne, T. Consistent strategy updating in spatial and non-spatial behavioral experiments does not promote cooperation in social networks. *PLoS One* **7**, e47718 (2012).
47. Gracia-Lázaro, C. *et al.* Heterogeneous networks do not promote cooperation when humans play a prisoner's dilemma. *Proc. Natl. Acad. Sci. USA* **109**, 12922–12926 (2012).
48. Gracia-Lázaro, C., Cuesta, J. A., Sanchez, A. & Moreno, Y. Human behavior in prisoner's dilemma experiments suppresses network reciprocity. *Sci. Rep.* **2** (2012).
49. Semmann, D. Conditional cooperation can hinder network reciprocity. *Proc. Natl. Acad. Sci. USA* **109**, 12846–12847 (2012).
50. Grujić, J. *et al.* A comparative analysis of spatial prisoner's dilemma experiments: Conditional cooperation and payoff irrelevance. *Sci. Rep.* **4**, 4615 (2014).
51. Rand, D. G., Nowak, M. A., Fowler, J. H. & Christakis, N. A. Static network structure can stabilize human cooperation. *Proc. Natl. Acad. Sci. USA* **111**, 17093–17098 (2014).
52. Li, X. *et al.* Punishment diminishes the benefits of network reciprocity in social dilemma experiments. *Proc. Natl. Acad. Sci. USA* **115**, 30–35 (2018).
53. Roca, C. P., Cuesta, J. A. & Sánchez, A. Promotion of cooperation on networks? the myopic best response case. *Eur. Phys. J. B* **71**, 587 (2009).
54. Orbell, J. M., Schwartz-Shea, P. & Simmons, R. T. Do cooperators exit more readily than defectors? *Am. Polit. Sci. Rev.* **78**, 147–162 (1984).
55. Cassar, A. Coordination and cooperation in local, random and small world networks: Experimental evidence. *Games Econ. Behav.* **58**, 209–230 (2007).
56. Kirchkamp, O. & Nagel, R. Naive learning and cooperation in network experiments. *Games Econ. Behav.* **58**, 269–292 (2007).
57. Traulsen, A., Semmann, D., Sommerfeld, R. D., Krambeck, H. J. & Milinski, M. Human strategy updating in evolutionary games. *Proc. Natl. Acad. Sci. USA* **107**, 2962–2966 (2010).
58. Grujić, J., Eke, B., Cabrales, A., Cuesta, J. A. & Sanchez, A. Three is a crowd in iterated prisoner's dilemmas: experimental evidence on reciprocal behavior. *Sci. Rep.* **2**, 638 (2012).
59. Amaral, M. A., Wardil, L., Perc, M. & da Silva, J. K. Stochastic win-stay-lose-shift strategy with dynamic aspirations in evolutionary social dilemmas. *Physical Review E* **94**, 032317 (2016).
60. Armano, G. & Javarone, M. A. The beneficial role of mobility for the emergence of innovation. *Scientific Reports* **7**, 1781 (2017).
61. Amaral, M. A. & Javarone, M. A. Heterogeneous update mechanisms in evolutionary games: mixing innovative and imitative dynamics. *Physical Review E* **97**, 042305 (2018).
62. Danku, Z., Wang, Z. & Szolnoki, A. Imitate or innovate: Competition of strategy updating attitudes in spatial social dilemma games. *EPL (Europhysics Letters)* **121**, 18002 (2018).
63. Hauert, C., De Monte, S., Hofbauer, J. & Sigmund, K. Volunteering as Red Queen mechanism for cooperation in public goods games. *Science* **296**, 1129–1132 (2002).
64. Szabó, G. & Hauert, C. Evolutionary prisoner's dilemma games with voluntary participation. *Phys. Rev. E* **66**, 062903 (2002).
65. Cardinot, M., Griffith, J. & O'Riordan, C. A further analysis of the role of heterogeneity in coevolutionary spatial games. *Physica A: Statistical Mechanics and its Applications* **493**, 116–124 (2018).

66. Cardinot, M., Griffith, J., O’Riordan, C. & Perc, M. Cooperation in the spatial prisoner’s dilemma game with probabilistic abstention. *Scientific Reports* **8** (2018).
67. van Veelen, M., García, J., Rand, D. G. & Nowak, M. A. Direct reciprocity in structured populations. *Proc. Natl. Acad. Sci. USA* **109**, 9929–9934 (2012).
68. Pacheco, J. M., Traulsen, A. & Nowak, M. A. Coevolution of strategy and structure in complex networks with dynamical linking. *Phys. Rev. Lett.* **97**, 258103 (2006).
69. Huang, K., Zheng, X., Li, Z. & Yang, Y. Understanding cooperative behavior based on the coevolution of game strategy and link weight. *Scientific Reports* **5**, 14783 (2015).
70. Fehl, K., van der Post, D. J. & Semmann, D. Co-evolution of behaviour and social network structure promotes human cooperation. *Ecol. Lett.* **14**, 546–551 (2009).
71. Rand, D. G., Arbesman, S. & Christakis, N. A. Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci. USA* **108**, 19193–19198 (2011).
72. Cuesta, J. A., Gracia-Lazaro, C., Ferrer, A., Moreno, Y. & Sanchez, A. Reputation drives cooperative behaviour and network formation in human groups. *Sci. Rep.* **5** (2015).
73. Perc, M. & Szolnoki, A. Coevolutionary games—a mini review. *BioSystems* **99**, 109–125 (2010).
74. Melamed, D., Harrell, A. & Simpson, B. Cooperation, clustering, and assortative mixing in dynamic networks. *Proc. Natl. Acad. Sci. USA* **115**, 951–956 (2018).
75. Grilo, C. & Luís, C. Effects of asynchronism on evolutionary games. *J. Theor. Biol.* **109** (2012).

Author Contributions

All authors were involved in the design of the research and in the analysis. F.D.R. performed the numerical analysis and produced the graphics; C.P. performed the network analysis; F.D. wrote the paper.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-41547-w>.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019