



# Indirect reciprocity with private, noisy, and incomplete information

Christian Hilbe<sup>a,1</sup>, Laura Schmid<sup>a</sup>, Josef Tkadlec<sup>a</sup>, Krishnendu Chatterjee<sup>a</sup>, and Martin A. Nowak<sup>b,c,d</sup>

<sup>a</sup>Institute of Science and Technology Austria, 3400 Klosterneuburg, Austria; <sup>b</sup>Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138; <sup>c</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; and <sup>d</sup>Department of Mathematics, Harvard University, Cambridge, MA 02138

Edited by Brian Skyrms, University of California, Irvine, CA, and approved October 16, 2018 (received for review June 19, 2018)

**Indirect reciprocity is a mechanism for cooperation based on shared moral systems and individual reputations. It assumes that members of a community routinely observe and assess each other and that they use this information to decide who is good or bad, and who deserves cooperation. When information is transmitted publicly, such that all community members agree on each other's reputation, previous research has highlighted eight crucial moral systems. These "leading-eight" strategies can maintain cooperation and resist invasion by defectors. However, in real populations individuals often hold their own private views of others. Once two individuals disagree about their opinion of some third party, they may also see its subsequent actions in a different light. Their opinions may further diverge over time. Herein, we explore indirect reciprocity when information transmission is private and noisy. We find that in the presence of perception errors, most leading-eight strategies cease to be stable. Even if a leading-eight strategy evolves, cooperation rates may drop considerably when errors are common. Our research highlights the role of reliable information and synchronized reputations to maintain stable moral systems.**

cooperation | indirect reciprocity | social norms | evolutionary game theory

**H**umans treat their reputations as a form of social capital (1–3). They strategically invest into their good reputation when their benevolent actions are widely observed (4–6), which in turn makes them more likely to receive benefits in subsequent interactions (7–12). Reputations undergo constant changes in time. They are affected by rumors and gossip (13), which themselves can spread in a population and develop a life of their own. Evolutionary game theory explores how good reputations are acquired and how they affect subsequent behaviors, using the framework of indirect reciprocity (14–17). This framework assumes that members of a population routinely observe and assess each other's social interactions. Whether a given action is perceived as good depends on the action itself, the context, and the social norm used by the population. Behaviors that yield a good reputation in one society may be condemned in others. A crucial question thus becomes: Which social norms are most conducive to maintain cooperation in a population?

Different social norms can be ordered according to their complexity (18) and according to the information that is required to assess a given action (19, 20). According to "first-order norms," the interpretation of an action depends only on the action itself. When a donor interacts with a recipient in a social dilemma, the donor's reputation improves if she cooperates, whereas her reputation drops if she defects (21–26). According to "second-order norms," the interpretation of an action additionally depends on the reputation of the recipient. The recipient's reputation provides the context of the interaction. It allows observers to distinguish between justified and unjustified defections (27–29). For example, the standing strategy considers it wrongful only to defect against well-reputed recipients; donors who defect against bad recipients do not suffer from

an impaired reputation (30). According to "third-order norms," observers need to additionally take the donor's reputation into account. In this way, assessment rules of higher order are increasingly able to give a more nuanced interpretation of a donor's action, but they also require observers to store and process more information.

When subjects are restricted to binary norms, such that reputations are either "good" or "bad," an exhaustive search shows there are eight third-order norms that maintain cooperation (20, 31). These "leading-eight strategies" are summarized in Table 1, and we refer to them as L1–L8. None of them is exclusively based on first-order information, whereas two of them (called "simple standing" and "stern judging," refs. 32 and 33) require second-order information only. There are several universal characteristics that all leading-eight strategies share. For example, against a recipient with a good reputation, a donor who cooperates should always obtain a good reputation, whereas a donor who defects should gain a bad reputation. The norms differ, however, in how they assess actions toward bad recipients. Whereas some norms allow good donors to preserve their good standing when they cooperate with a bad recipient, other norms disincentivize such behaviors.

Ohtsuki and Iwasa (20, 31) have shown that if all members of a population adopt a leading-eight strategy, stable cooperation can emerge. Their model, however, assumes that the players' images are synchronized; two population members would always agree on the current reputation of some third population member. The assumption of publicly available and synchronized information

## Significance

**Indirect reciprocity explores how humans act when their reputation is at stake, and which social norms they use to assess the actions of others. A crucial question in indirect reciprocity is which social norms can maintain stable cooperation in a society. Past research has highlighted eight such norms, called "leading-eight" strategies. This past research, however, is based on the assumption that all relevant information about other population members is publicly available and that everyone agrees on who is good or bad. Instead, here we explore the reputation dynamics when information is private and noisy. We show that under these conditions, most leading-eight strategies fail to evolve. Those leading-eight strategies that do evolve are unable to sustain full cooperation.**

Author contributions: C.H., L.S., J.T., K.C., and M.A.N. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>To whom correspondence should be addressed. Email: christian.hilbe@ist.ac.at.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810565115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1810565115/-DCSupplemental).

Published online November 14, 2018.

**Table 1. The leading-eight strategies of indirect reciprocity**

Assessment rule	<div style="display: flex; justify-content: space-around; font-size: small;"> <span>*</span> <span>Consistent Standing</span> <span>Simple Standing</span> <span>*</span> <span>*</span> <span>Stern Judging</span> <span>Staying</span> <span>Judging</span> </div>							
	L1	L2	L3	L4	L5	L6	L7	L8
Good cooperates with Good	g	g	g	g	g	g	g	g
Good cooperates with Bad	g	b	g	g	b	b	g	b
Bad cooperates with Good	g	g	g	g	g	g	g	g
Bad cooperates with Bad	g	g	g	b	g	b	b	b
Good defects against Good	b	b	b	b	b	b	b	b
Good defects against Bad	g	g	g	g	g	g	g	g
Bad defects against Good	b	b	b	b	b	b	b	b
Bad defects against Bad	b	b	g	g	g	g	b	b
Action rule	L1	L2	L3	L4	L5	L6	L7	L8
Good meets Good	C	C	C	C	C	C	C	C
Good meets Bad	D	D	D	D	D	D	D	D
Bad meets Good	C	C	C	C	C	C	C	C
Bad meets Bad	C	C	D	D	D	D	D	D

There are eight strategies, called the “leading eight,” that have been shown to maintain cooperation under public assessment (20, 31). Each such strategy consists of an assessment rule and of an action rule. The assessment rule determines whether a donor is deemed good (g) or bad (b). This assessment depends on the context of the interaction (on the reputations of the donor and the recipient) and on the donor’s action (C or D). The action rule determines whether to cooperate with a given recipient when in the role of the donor. A donor’s action may depend on her own reputation, as well as on the reputation of the recipient. All of the leading-eight strategies agree that cooperation against a good player should be deemed as good, whereas defection against a good player should be deemed bad. They disagree in how they evaluate actions toward bad recipients.

greatly facilitates a rigorous analysis of the reputation dynamics. Yet in most real populations, different individuals may have access to different kinds of information, and thus they might disagree on how they assess others. Their opinions may well be correlated, but they will not be correlated perfectly. Once individuals disagree in their initial evaluation of some person, their views may further diverge over time. How such initial disagreements spread may itself depend on the social norm used by the population. While some norms can maintain cooperation even in the presence of rare disagreements, other norms are more susceptible to deviations from the public information assumption (34–37). Here, we explore systematically how the leading-eight strategies fare when information is private, noisy, and incomplete. We show that under these conditions, most leading-eight strategies cease to be stable. Even if a leading-eight strategy evolves, the resulting cooperation rate may be drastically reduced.

**Results**

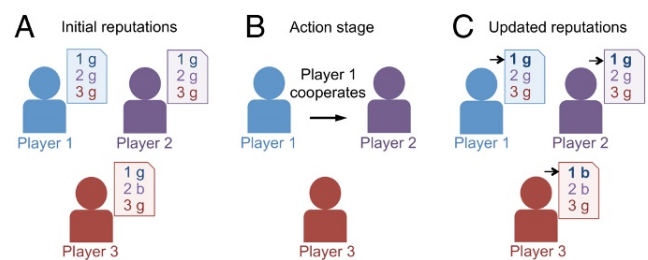
**Model Setup.** We consider a well-mixed population of size  $N$ . The members of this population are engaged in a series of cooperative interactions. In each round, two individuals are randomly drawn, a donor and a recipient. The donor can then decide whether to transfer a benefit  $b$  to the recipient at own cost  $c$ , with  $0 < c < b$ . We refer to the donor’s two possible actions as cooperation (transferring the benefit) and defection (not doing anything). Whereas the donor and the recipient always learn the donor’s decision, each other population member independently learns the donor’s decision with probability  $q > 0$ . Observations may be subject to noise: We assume that all players who learn the donor’s action may misperceive it with probability  $\epsilon > 0$ , independently of the other players. In that case,

a player misinterprets the donor’s cooperation as defection or, conversely, the donor’s defection as cooperation. After observing an interaction, population members independently update their image of the donor according to the information they have (Fig. 1).

To do so, we assume that each individual is equipped with a strategy that consists of an assessment rule and an action rule. The player’s assessment rule governs how players update the reputation they assign to the donor. Here we consider third-order assessment rules. That is, when updating the donor’s reputation, a player takes the donor’s action into account, as well as the donor’s and the recipient’s previous reputation. Importantly, when two observers differ in their initial assessment of a given donor, they may also disagree on the donor’s updated reputation, even if both apply the same assessment rule and observe the same interaction (Fig. 1C). The second component of a player’s strategy, the action rule, determines which action to take when chosen to be the donor. This action may depend on the player’s own reputation, as well as on the reputation of the recipient. A player’s payoff for this indirect reciprocity game is defined as the expected benefit obtained as a recipient, reduced by the expected costs paid when acting as a donor, averaged over many rounds (see *Materials and Methods* for details).

**Analysis of the Reputation Dynamics.** We first explore how different social norms affect the dynamics of reputations, keeping the strategies of all players fixed. To this end, we use the concept of image matrices (34–36). These matrices record, at any point in time, which reputations players assign to each other. In Fig. 2 *A–H*, we show a snapshot of these image matrices for eight different scenarios. In all scenarios, the population consists in equal proportions of a leading-eight strategy, of unconditional cooperators who regard everyone as good (ALLC) and of unconditional defectors who regard everyone as bad (ALLD). Depending on the leading-eight strategy considered, the reputation dynamics in these scenarios can differ considerably.

First, for four of the eight scenarios, a substantial proportion of leading-eight players assigns a good reputation to ALLD players. The average proportion of ALLD players considered



**Fig. 1.** Under indirect reciprocity, individual actions are continually assessed by all population members. (A) We consider a population of different players. All players hold a private repository where they store which of their coplayers they deem as either good (g) or bad (b). Different players may hold different views on the same coplayer. In this example, player 2 is considered to be good from the perspective of the first two players, but he is considered to be bad by player 3. (B) In the action stage, two players are randomly chosen, a donor (here, player 1) and a recipient (here, player 2). The donor can then decide whether or not to cooperate with the recipient. The donor’s decision may depend on the stored reputations in her own private repository. (C) After the action stage, all players who observe the interaction update the donor’s reputation. The newly assigned reputation may differ across the population even if all players apply the same social norm. This can occur (i) when individuals already disagreed on their initial assessments of the involved players, (ii) when some subjects do not observe the interaction and hence do not update the donor’s reputation accordingly, or (iii) when there are perception errors.

as good by L3, L4, L5, and L6 is given by 31%, 31%, 42%, and 50%, respectively (*SI Appendix, Fig. S1*). In terms of these four leading-eight strategies, a bad player who defects against another bad player deserves a good reputation (Table 1). In particular, ALLD players can easily gain a good reputation whenever they encounter another ALLD player. Moreover, the higher the proportion of ALLD players in a population, the more readily they obtain a good reputation. This finding suggests that while L3–L6 might be stable when these strategies are common in the population (20, 38), they have problems in restraining the payoff of ALLD when defectors are predominant.

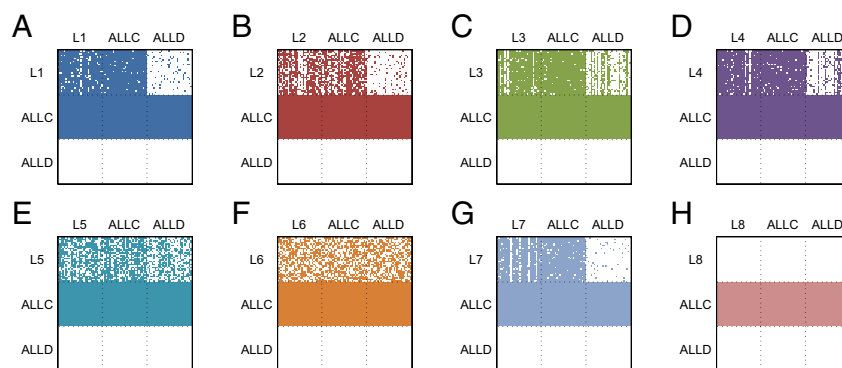
Second, leading-eight players may sometimes collectively judge a player of their own kind as bad. In Fig. 2, such cases are represented by white vertical lines in the upper left square of an image matrix. In *SI Appendix, Fig. S2* we show that such apparent misjudgments are typically introduced by perception errors. They occur, for example, when a leading-eight donor defects against an ALLC recipient, who is mistakenly considered as bad by the donor. Other leading-eight players who witness this interaction will then collectively assign a bad reputation to the donor—in their eyes, a good recipient has not obtained the help he deserves. This example highlights that under private information, an isolated disagreement about the reputation of some population member can have considerable consequences on the further reputation dynamics.

To gain a better understanding of such cases, we analytically explored the consequences of a single disagreement in a homogeneous population of leading-eight players (see *SI Appendix* for all details). There we assume that initially, all players consider each other as good, with the exception of one player who considers a random coplayer as bad. Assuming that no further errors occur, we study how likely the population recovers from this single disagreement (i.e., how likely the population reverts to a state where everyone is considered good) and how long it takes until recovery. While some leading-eight strategies are guaranteed to recover from single disagreements, we find that other strategies may reach an absorbing state where players mutually assign a bad reputation to each other. Moreover, even if recovery occurs, for some strategies it may take a considerable time (*SI Appendix, Fig. S3*). Two strategies fare particularly badly: L6 and L8 have the lowest probability to recover from a

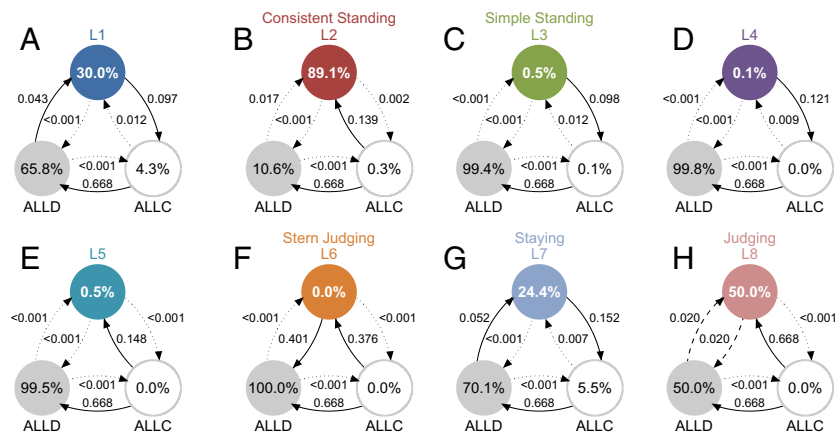
single disagreement, and they have the longest recovery time. This finding is also reflected in Fig. 2, which shows that these two strategies are unable to maintain cooperation. L6 eventually assigns random reputations to all coplayers, whereas L8 assigns a bad reputation to everyone (*SI Appendix, Fig. S4*). While L6 (“stern”) has been found to be particularly successful under public information (18, 32, 33), our results confirm that this strategy is too strict and unforgiving when information is private and noisy (34–36).

**Evolutionary Dynamics.** Next we explore how likely a leading-eight strategy would evolve when population members can change their strategies over time. We first consider a minimalistic scenario, where players can choose among three strategies only, a leading-eight strategy  $L_i$ , ALLC, and ALLD. To model how players adopt new strategies, we consider simple imitation dynamics (39–42). In each time step of the evolutionary process, one player is picked at random. With probability  $\mu$  (the mutation rate), this player then adopts some random strategy, corresponding to the case of undirected learning. With the remaining probability  $1 - \mu$ , the player randomly chooses a role model from the population. The higher the payoff of the role model, the more likely it is that the focal player adopts the role model’s strategy (*Materials and Methods*). Overall, the two modes of updating, mutation and imitation, give rise to an ergodic process on the space of all population compositions. In the following, we present results for the case when mutations are relatively rare (43, 44).

First, we calculated for a fixed benefit-to-cost ratio of  $b/c = 5$  how often each strategy is played over the course of evolution, for each of the eight possible scenarios (Fig. 3). In four cases, the leading-eight strategy is played in less than 1% of the time. These cases correspond to the four leading-eight strategies L3–L6 that frequently assign a good reputation to ALLD players. For these leading-eight strategies, once everyone in a population has learned to be a defector, players have difficulties in reestablishing a cooperative regime (in Fig. 3 C–F, once ALLD is reached, every other strategy has a fixation probability smaller than 0.001). In contrast, the strategy L8 is played in substantial proportions. But in the presence of noise, players with this strategy always defect, because they deem everyone as bad (Fig. 2).



**Fig. 2.** (A–H) When individuals base their decisions on noisy private information, their assessments may diverge. Models of private information need to keep track of which player assigns which reputation to which coplayer at any given time. These pairwise assessments are represented by image matrices. Here, we represent these image matrices graphically, assuming that the population consists of equal parts of a leading-eight strategy, of unconditional cooperators (ALLC) and unconditional defectors (ALLD). A colored dot means that the corresponding row player assigns a good reputation to the column player. Without loss of generality, we assume that ALLC players assign a good reputation to everyone, whereas ALLD players deem everyone as bad. The assessments of the leading-eight players depend on the coplayer’s strategy and on the frequency of perception errors. We observe that two of the leading-eight strategies are particularly prone to errors: L6 (“stern judging”) eventually assigns a random reputation to any coplayer, while L8 (“judging”) eventually considers everyone as bad. Only the other six strategies separate between conditionally cooperative strategies and unconditional defectors. Each box shows the image matrix after  $2 \cdot 10^6$  simulated interactions in a population of size  $N = 3 \cdot 30 = 90$ . Perception errors occur at rate  $\varepsilon = 0.05$ , and interactions are observed with high probability,  $q = 0.9$ .



**Fig. 3.** Most of the leading-eight strategies are disfavored in the presence of perception errors. We simulated the evolutionary dynamics when each of the leading-eight strategies competes with ALLC and ALLD. These simulations assume that, over time, players tend to imitate coplayers with more profitable strategies and that they occasionally explore random strategies (*Materials and Methods*). The numbers within the circles represent the abundance of the respective strategy in the selection–mutation equilibrium. The numbers close to the arrows represent the fixation probability of a single mutant into the given resident strategy. We use solid lines for the arrows to depict a fixation probability that exceeds the neutral probability  $1/N$ , and we use dotted lines if the fixation probability is smaller than  $1/N$ . In four cases, we find that ALLD is predominant (C–F). In one case (H), the leading-eight strategy coexists with ALLD but without any cooperation. In the remaining cases (A, B, and G), we find that L1 and L7 are played with moderate frequencies, but only populations that have access to L2 (“consistent standing”) settle at the leading-eight strategy. Parameters: Population size  $N = 50$ , benefit  $b = 5$ , cost  $c = 1$ , strength of selection  $s = 1$ , error rate  $\varepsilon = 0.05$ , observation probability  $q = 0.9$ , in the limit of rare mutations  $\mu \rightarrow 0$ .

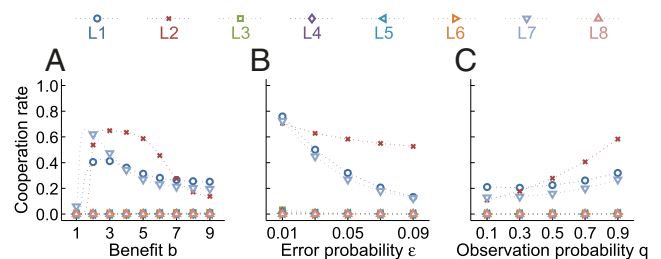
There are only three scenarios in Fig. 3 that allow for positive cooperation rates. The corresponding leading-eight strategies are L1, L2 (“consistent standing”), and L7 (“staying,” ref. 45). For L1 and L7, the evolutionary dynamics take the form of a rock–scissors–paper cycle (46–50). The leading-eight strategy can be invaded by ALLC, which gives rise to ALLD, which in turn leads back to the leading-eight strategy. Because ALLD is most robust in this cycle, the leading-eight strategies are played in less than one-third of the time (Fig. 3A and G).

Only consistent standing, L2, is able to compete with ALLC and ALLD in a direct comparison (Fig. 3B). Under consistent standing, there is a unique action in each possible situation that allows a donor to obtain a good standing. For example, when a good donor meets a bad recipient, the donor keeps her good standing by defecting, but loses it by cooperating. Compared with stern judging, which has a similar property (18), consistent standing incentivizes cooperation more strongly. When two bad players interact, the correct decision according to consistent standing is to cooperate, whereas a stern player would defect (Table 1).

Nevertheless, we find that even when consistent standing is common, the average cooperation rate in the population rarely exceeds 65%. To show this, we repeated the previous evolutionary simulations for the eight scenarios while varying the benefit-to-cost ratio, the error rate, and the observation probability (Fig. 4). These simulations confirm that five of the leading-eight strategies cannot maintain any cooperation when competing with ALLC and ALLD. Only for L1, L2, and L7 are average cooperation rates positive, reaching a maximum for intermediate benefit-to-cost ratios (Fig. 4A). If the benefit-to-cost ratio is too low, we find that each of these leading-eight strategies can be invaded by ALLD, whereas if the ratio is too high, ALLC can invade (SI Appendix, Fig. S5). In between, consistent standing may outperform ALLC and ALLD, but in the presence of noise it does not yield high cooperation rates against itself. Even if all interactions are observed ( $q = 1$ ), cooperation rates in a homogeneous L2 population drop below 70% once the error rate exceeds 5% (SI Appendix, Fig. S4). Our analytical results in SI Appendix suggest that while L2 populations always recover from single disagreements, it may take them a substantial time to do so, during

which further errors may accumulate. As a result, whereas L2 seems most robust when coevolving with ALLC and ALLD, it is unable to maintain full cooperation. Furthermore, additional simulation results suggest that even if L2 is able to resist invasion by ALLC and ALLD, it may be invaded by mutant strategies that differ in only one bit from L2 (SI Appendix, Fig. S6).

So far, we have assumed that mutations are rare, such that populations are typically homogeneous. Experimental evidence, however, suggests that there is considerable variation in the social norms used by subjects (4, 7–11). While some subjects are best classified as unconditional defectors, others act as unconditional cooperators or use more sophisticated higher-order strategies (11). In agreement with these experimental studies, there is theoretical evidence that some leading-eight strategies like L7 may form stable coexistences with ALLC (36). In SI Appendix, Figs. S7–S9, we present further evolutionary results for higher mutation rates, in which such coexistences are possible.



**Fig. 4.** Noise can prevent the evolution of full cooperation even if leading-eight strategies evolve. We repeated the evolutionary simulations in Fig. 3, but varying (A) the benefit of cooperation, (B) the error rate, and (C) the observation probability. The graph shows the average cooperation rate for each scenario in the selection–mutation equilibrium. This cooperation rate depends on how abundant each strategy is in equilibrium and on how much cooperation each strategy yields against itself in the presence of noise. For five of the eight scenarios, cooperation rates remain low across the considered parameter range. Only the three other leading-eight strategies can persist in the population, but even then cooperation rates typically remain below 70%. We use the same baseline parameters as in Fig. 3.

There we show that in the three cases L1, L2, and L7, populations may consist of a mixture of the leading-eight strategy and ALLC for a considerable time. However, in agreement with our rare-mutation results, we find for L1 and L7 that this mixture of leading-eight strategy and ALLC is susceptible to stochastic invasion by ALLD.

### Discussion

Indirect reciprocity explores how cooperation can be maintained when individuals assess and act on each other's reputations. Simple strategies of indirect reciprocity like image scoring (21, 22) have been suspected to be unstable, because players may abstain from punishing defectors to maintain their own good score (27). In contrast, the leading-eight strategies additionally take the context of an interaction into account. They have been considered to be prime candidates for stable norms that maintain cooperation (20, 31). Corresponding models, however, assume that each pairwise interaction is witnessed only by one observer, who disseminates the outcome of the interaction to all other population members. As a consequence, the resulting opinions within a population will be perfectly synchronized. Even if donors are subject to implementation errors, or if the observer misperceives an interaction, all players will have the same image of the donor after the interaction has taken place.

While the assumption of perfectly synchronized reputations is a useful idealization, we believe that it may be too strict in some applications. Subjects often differ in the prior information they have, and even if everyone has access to the same information [as is often the case in online platforms (51, 52)], individuals differ in how much weight they attribute to different pieces of evidence. As a result, individuals might disagree on each other's reputations. These disagreements can proliferate over time. Herein, we have thus systematically compared the performance of the leading-eight strategies when information is incomplete, private, and noisy. The leading-eight strategies differ in how they are affected by the noise introduced by private perception errors. Strategies like stern judging, that have been shown to be highly successful under public information (18, 32, 33), fail to distinguish between friend and foe when information is private. While we have considered well-mixed populations in which all players are connected, this effect might be even more pronounced when games take place on a network (53, 54). If players are able only to observe interactions between players in their immediate neighborhood, network-structured populations may amplify the problem of incomplete information. Pairwise interactions that one player is able to observe may be systematically hidden from his neighbor's view. Thus, the study of indirect reciprocity on networks points to an interesting direction for future research.

The individuals in our model are completely independent when forming their beliefs. In particular, they are not affected by the opinions of others, swayed by gossip and rumors, or engaged in communication. Experimental evidence suggests that even when all subjects witness the same social interaction, gossip can greatly modify beliefs and align the subjects' subsequent behaviors (13). Seen from this angle, our study highlights the importance of coordination and communication for the stability of indirect reciprocity. Social norms that fail when information is noisy and private may sustain full cooperation when information is mutually shared and discussed.

### Materials and Methods

**Model Setup.** We consider  $N$  individuals in a well-mixed population. Each player's strategy is given by a pair  $(\alpha, \beta)$ . The first component,

$$\alpha = (\alpha_{gCg}, \alpha_{gCb}, \alpha_{bCg}, \alpha_{bCb}, \alpha_{gDg}, \alpha_{gDb}, \alpha_{bDg}, \alpha_{bDb}), \quad [1]$$

corresponds to the player's assessment rule. An entry  $\alpha_{xAy}$  is equal to one if the player assigns a good reputation to a donor of reputation  $x$  who chooses action  $A$  against a recipient with reputation  $y$ . Otherwise, if such a donor is considered as bad, the corresponding entry is zero. The second component of the strategy,

$$\beta = (\beta_{gg}, \beta_{gb}, \beta_{bg}, \beta_{bb}), \quad [2]$$

gives the player's action rule. An entry  $\beta_{xy}$  is equal to one if the focal player with reputation  $x$  cooperates with a recipient with reputation  $y$ ; otherwise it is zero. The assessment and action rules of the leading-eight strategies are shown in Table 1. We define ALLC as the strategy with assessment rule  $\alpha = (1, \dots, 1)$  and action rule  $\beta = (1, \dots, 1)$ . ALLD is the strategy with  $\alpha = (0, \dots, 0)$  and  $\beta = (0, \dots, 0)$ .

**Reputation Dynamics.** To simulate the reputation dynamics for players with fixed strategies, we consider the image matrix (34–36)  $M(t) = (m_{ij}(t))$  of a population at time  $t$ . Its entries satisfy  $m_{ij}(t) = 1$  if player  $i$  deems player  $j$  as good at time  $t$  and  $m_{ij}(t) = 0$  otherwise. We assume that initially, all players have a good reputation,  $m_{ij}(0) = 1$  for all  $i, j$ . However, our results are unchanged if the players' initial reputations are assigned randomly. We get only slightly different results if all initial reputations are bad; in that case, L7 players are unable to acquire a good reputation over the course of the game (for details, see *SI Appendix*).

In each round  $t$ , two players  $i$  and  $j$  are drawn from the population at random, a donor and a recipient. The donor then decides whether to cooperate. Her choice is uniquely determined by her action rule  $\beta$  and by the reputations she assigns to herself and to the recipient,  $m_{ii}(t)$  and  $m_{ij}(t)$ . The donor and the recipient always observe the donor's decision; all other players independently observe it with probability  $q$ . With probability  $\varepsilon$ , a player who observes the donor's action misperceives it, independent of the other players. All players who observe the interaction update their assessment of the donor according to their assessment rule. This yields the image matrix  $M(t + 1)$ .

We iterate the above elementary process over many rounds (our numbers are based on  $10^6$  rounds or more). Based on these simulations, we can now calculate how often player  $i$  considers  $j$  to be good on average and how often player  $i$  cooperates with  $j$  on average. If the estimated pairwise cooperation rate of  $i$  against  $j$  is given by  $\hat{x}_{ij}$ , we define player  $i$ 's payoff as  $\hat{\pi}_i = \frac{1}{N-1} \sum_{j \neq i} b\hat{x}_{ji} - c\hat{x}_{ij}$ .

**Evolutionary Dynamics.** On a larger timescale, we assume that players can change their strategies  $(\alpha, \beta)$ . To model the strategy dynamics, we consider a pairwise comparison process (39–41). In each time step of this process, one individual is randomly chosen from the population. With probability  $\mu$  this individual then adopts a random strategy, with all other available strategies having the same probability to be picked. With the remaining probability  $1 - \mu$  the focal individual  $i$  chooses a random role model  $j$  from the population. If the players' payoffs are  $\hat{\pi}_i$  and  $\hat{\pi}_j$ , player  $i$  adopts  $j$ 's strategy with probability  $P(\hat{\pi}_j, \hat{\pi}_i) = (1 + \exp[-s(\hat{\pi}_j - \hat{\pi}_i)])^{-1}$  (55). The parameter  $s \geq 0$  is the "strength of selection." It measures how strongly imitation events are biased in favor of strategies with higher payoffs. For  $s = 0$  we obtain  $P(\hat{\pi}_j, \hat{\pi}_i) = 1/2$ , and imitation occurs at random. As  $s$  increases, payoffs become increasingly relevant when  $i$  considers imitating  $j$ 's strategy.

In the main text, we assume players can choose only between a leading-eight strategy  $L_i$ , ALLC, and ALLD. As we show in *SI Appendix, Fig. S6*, the stability of a leading-eight strategy may be further undermined if additional mutant strategies are available. Moreover, in the main text we report only results when mutations are comparably rare (43, 44). In *SI Appendix, Figs. S7–S9* we show further results for substantial mutation rates. Given the players' payoffs for each possible population composition, the selection–mutation equilibrium can be calculated explicitly. All details are provided in *SI Appendix*.

**ACKNOWLEDGMENTS.** This work was supported by the European Research Council Start Grant 279307 Graph Games (to K.C.), Austrian Science Fund (FWF) Grant P23499-N23 (to K.C.), FWF Nationale Forschungsnetzwerke Grant S11407-N23 Rigorous Systems Engineering/Systematic Methods in Systems Engineering (to K.C.), Office of Naval Research Grant N00014-16-1-2914 (to M.A.N.), and the John Templeton Foundation (M.A.N.). C.H. acknowledges generous support from the ISTFELLOW program.

1. Alexander R (1987) *The Biology of Moral Systems* (Aldine de Gruyter, New York).
2. Rand DG, Nowak MA (2012) Human cooperation. *Trends Cogn Sci* 117:413–425.
3. Melis AP, Semmann D (2010) How is human cooperation different? *Philos Trans R Soc B* 365:2663–2674.

4. Engelmann D, Fischbacher U (2009) Indirect reciprocity and strategic reputation building in an experimental helping game. *Games Econ Behav* 67:399–407.
5. Jacquet J, Hauert C, Traulsen A, Milinski M (2011) Shame and honour drive cooperation. *Biol Lett* 7:899–901.

6. Ohtsuki H, Iwasa Y, Nowak MA (2015) Reputation effects in public and private interactions. *PLoS Comput Biol* 11:e1004527.
7. Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. *Science* 288:850–852.
8. Seinen I, Schram A (2006) Social status and group norms: Indirect reciprocity in a repeated helping experiment. *Eur Econ Rev* 50:581–602.
9. Bolton G, Katok E, Ockenfels A (2005) Cooperation among strangers with limited information about reputation. *J Public Econ* 89:1457–1468.
10. van Apeldoorn J, Schram A (2016) Indirect reciprocity; a field experiment. *PLoS One* 11:e0152076.
11. Swakman V, Molleman L, Ule A, Egas M (2016) Reputation-based cooperation: Empirical evidence for behavioral strategies. *Evol Hum Behav* 37:230–235.
12. Capraro V, Giardini F, Vilone D, Paolucci M (2016) Partner selection supported by opaque reputation promotes cooperative behavior. *Judgment Decis Making* 11:589–600.
13. Sommerfeld RD, Krambeck HJ, Semmann D, Milinski M (2007) Gossip as an alternative for direct observation in games of indirect reciprocity. *Proc Natl Acad Sci USA* 104:17435–17440.
14. Nowak MA, Sigmund K (2005) Evolution of indirect reciprocity. *Nature* 437:1291–1298.
15. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314:1560–1563.
16. Sigmund K (2010) *The Calculus of Selfishness* (Princeton Univ Press, Princeton).
17. Sigmund K (2012) Moral assessment in indirect reciprocity. *J Theor Biol* 299:25–30.
18. Santos FP, Santos FC, Pacheco JM (2018) Social norm complexity and past reputations in the evolution of cooperation. *Nature* 555:242–245.
19. Brandt H, Sigmund K (2004) The logic of reprobation: Assessment and action rules for indirect reciprocation. *J Theor Biol* 231:475–486.
20. Ohtsuki H, Iwasa Y (2004) How should we define goodness?—Reputation dynamics in indirect reciprocity. *J Theor Biol* 231:107–120.
21. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393:573–577.
22. Nowak MA, Sigmund K (1998) The dynamics of indirect reciprocity. *J Theor Biol* 194:561–574.
23. Ohtsuki H (2004) Reactive strategies in indirect reciprocity. *J Theor Biol* 227:299–314.
24. Brandt H, Sigmund K (2005) Indirect reciprocity, image scoring, and moral hazard. *Proc Natl Acad Sci USA* 102:2666–2670.
25. Berger U (2011) Learning to cooperate via indirect reciprocity. *Games Econ Behav* 72:30–37.
26. Berger U, Grüne A (2016) On the stability of cooperation under indirect reciprocity with first-order information. *Games Econ Behav* 98:19–33.
27. Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. *Proc R Soc B* 268:745–753.
28. Panchanathan K, Boyd R (2003) A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *J Theor Biol* 224:115–126.
29. Suzuki S, Akiyama E (2007) Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity. *J Theor Biol* 245:539–552.
30. Sugden R (1986) *The Economics of Rights, Co-Operation and Welfare* (Blackwell, Oxford).
31. Ohtsuki H, Iwasa Y (2006) The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *J Theor Biol* 239:435–444.
32. Chalub FACC, Santos FC, Pacheco JM (2006) The evolution of norms. *J Theor Biol* 241:233–240.
33. Santos FP, Santos FC, Pacheco JM (2016) Social norms of cooperation in small-scale societies. *PLoS Comput Biol* 12:e1004709.
34. Uchida S (2010) Effect of private information on indirect reciprocity. *Phys Rev E* 82:036111.
35. Uchida S, Sasaki T (2013) Effect of assessment error and private information on stern-judging in indirect reciprocity. *Chaos Solitons Fractals* 56:175–180.
36. Okada I, Sasaki T, Nakai Y (2017) Tolerant indirect reciprocity can boost social welfare through solidarity with unconditional cooperators in private monitoring. *Sci Rep* 7:9737.
37. Okada I, Sasaki T, Nakai Y (2018) A solution for private assessment in indirect reciprocity using solitary observation. *J Theor Biol* 455:7–15.
38. Martinez-Vaquero LA, Cuesta JA (2013) Evolutionary stability and resistance to cheating in an indirect reciprocity model based on reputation. *Phys Rev E* 87:052810.
39. Traulsen A, Nowak MA, Pacheco JM (2006) Stochastic dynamics of invasion and fixation. *Phys Rev E* 74:011909.
40. Stewart AJ, Plotkin JB (2013) From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proc Natl Acad Sci USA* 110:15348–15353.
41. Reiter JG, Hilbe C, Rand DG, Chatterjee K, Nowak MA (2018) Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nat Commun* 9:555.
42. Hilbe C, Šimsa Š, Chatterjee K, Nowak MA (2018) Evolution of cooperation in stochastic games. *Nature* 559:246–249.
43. Fudenberg D, Imhof LA (2006) Imitation processes with small mutations. *J Econ Theor* 131:251–262.
44. Wu B, Gokhale CS, Wang L, Traulsen A (2012) How small are small mutation rates? *J Math Biol* 64:803–827.
45. Sasaki T, Okada I, Nakai Y (2017) The evolution of conditional moral assessment in indirect reciprocity. *Sci Rep* 7:41870.
46. Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics* (Cambridge Univ Press, Cambridge, UK).
47. Szolnoki A, Szabó G (2004) Phase transitions for rock-scissors-paper game on different networks. *Phys Rev E* 70:037102.
48. Claussen JC, Traulsen A (2008) Cyclic dominance and biodiversity in well-mixed populations. *Phys Rev Lett* 100:058104.
49. Szolnoki A, et al. (2014) Cyclic dominance in evolutionary games: A review. *J R Soc Interface* 11:20140735.
50. Stewart AJ, Parsons TL, Plotkin JB (2016) Evolutionary consequences of behavioral diversity. *Proc Natl Acad Sci USA* 113:E7003–E7009.
51. Resnick P, Zeckhauser R, Swanson J, Lockwood K (2006) The value of reputation on eBay: A controlled experiment. *Exp Econ* 9:79–101.
52. Restivo M, van de Rijt A (2012) Experimental study of informal rewards in peer production. *PLoS One* 7:e34358.
53. Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. *Nature* 433:312–316.
54. Szabó G, Fáth G (2007) Evolutionary games on graphs. *Phys Rep* 446:97–216.
55. Szabó G, Töke C (1998) Evolutionary prisoner's dilemma game on a square lattice. *Phys Rev E* 58:69–73.

# Supporting Information

## Indirect reciprocity with private, noisy and incomplete information

Christian Hilbe<sup>1</sup>, Laura Schmid<sup>1</sup>, Josef Tkadlec<sup>1</sup>, Krishnendu Chatterjee<sup>1</sup>, and Martin A. Nowak<sup>2</sup>

<sup>1</sup>IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

<sup>2</sup>Program for Evolutionary Dynamics, Department of Organismic and Evolutionary Biology and Department of Mathematics, Harvard University, Cambridge MA 02138, USA

We are interested in the evolution of indirect reciprocity under incomplete and noisy information. To this end, we describe the dynamics on two separate timescales. First, we consider the *game dynamics*. Here, we take the players' strategies as given, and we compute how the players' reputations change over time, how often they cooperate, and which payoffs they obtain on average. Second, we describe the *evolutionary dynamics*. Here, we allow players to change their strategies over time, assuming that strategies that yield a comparably high payoff are more likely to be adopted. In the following, we describe these two dynamics in more detail.

### 1 Game dynamics

#### 1.1 Description of the indirect reciprocity game

We consider a well-mixed population of size  $N$  whose members engage in a series of cooperative interactions. In each time step, two members of the population are randomly drawn from the population, a donor and a recipient. The donor can then decide whether or not to pay a cost  $c > 0$  to transfer a benefit  $b > c$  to the recipient. We interpret paying the cost as cooperation ( $C$ ) and refusing to do so as defection ( $D$ ). The donor's decision is partially observable: the donor and the receiver always learn whether or not the donor decided to cooperate, whereas the other population members independently observe it with probability  $q$ . Observations may be subject to noise: with probability  $\varepsilon$  the donor's action is misinterpreted such that a  $C$  is taken for a  $D$ , or vice versa. All players are equally likely to misperceive an action, independent of whether or not they actively took part in the respective interaction. We assume that there are infinitely many rounds in which players are asked to cooperate, and that the players' payoffs for the indirect reciprocity game are defined as their average payoffs over time (explicit definitions will be provided in the next section).

To make their cooperation decisions, we assume that each player is equipped with a strategy and a private reputation repository. The reputation repository is used to keep track of the reputations of all population members (as illustrated in **Fig. 1**). In line with a large body of the previous literature (1–3), we assume that reputations are binary: players are either considered as 'good' or 'bad'. We represent the

state of player  $i$ 's reputation repository at time  $t$  by an  $N$ -dimensional vector

$$\mathbf{m}_i(t) = (m_{i1}(t), \dots, m_{iN}(t)). \quad [1]$$

The entries  $m_{ij}$  of this vector are either 0 or 1, whereby  $m_{ij}(t) = 1$  means that player  $i$  assigns a good reputation to player  $j$  at time  $t$ , whereas  $m_{ij}(t) = 0$  means that  $i$  views  $j$  as bad. We assume that  $\mathbf{m}_i(0) = (1, \dots, 1)$  for all players  $i$ . That is, initially all players consider everyone as good.

A player's strategy then needs to tell the player ( $i$ ) how to update the entries in her reputation repository after observing a donor's action, and ( $ii$ ) whether to cooperate if the focal player finds herself in the role of the donor. The first aspect is determined by the player's assessment rule. Assessment rules take the form of an 8-dimensional vector

$$\alpha = (\alpha_{gCg}, \alpha_{gCb}, \alpha_{bCg}, \alpha_{bCb}, \alpha_{gDg}, \alpha_{gDb}, \alpha_{bDg}, \alpha_{bDb}). \quad [2]$$

The entries of  $\alpha$  can again take either the value 0 or 1. An entry  $\alpha_{xAy} = 1$  means that it is regarded as good if a donor with reputation  $x$  chooses action  $A$  against a recipient with reputation  $y$ . Analogously,  $\alpha_{xAy} = 0$  means that the focal player assigns a bad reputation to the corresponding donor. We note that after any interaction, players only update the donor's reputation; the reputation of the receiver remains unaffected. Assessment rules of the form [2] are called third-order assessment rules, since they depend on the donor's action, on the reputation of the donor, and on the reputation of the recipient. If the assessment is independent of the donor's current reputation, the assessment rule is referred to as second-order; and if the assessment is independent of both players' reputations, the assessment rule is referred to as first-order. The two unconditional cases  $\alpha = \mathbf{1} := (1, \dots, 1)$  and  $\alpha = \mathbf{0} := (0, \dots, 0)$ , according to which donors are always considered as good or bad, respectively, are zeroth-order assessment rules.

If the focal player finds herself in the role of the donor, her decision is determined by her action rule. Action rules are given by a 4-dimensional vector

$$\beta = (\beta_{gg}, \beta_{gb}, \beta_{bg}, \beta_{bb}). \quad [3]$$

An entry  $\beta_{xy} = 1$  indicates that a respective donor chooses to cooperate if her own reputation is  $x$  and the recipient's reputation is  $y$ . An entry  $\beta_{xy} = 0$  indicates that she defects. A player's strategy  $\sigma = (\alpha; \beta)$  is a combination of an assessment rule and an action rule. Since there are  $2^8$  different assessment rules and  $2^4$  different action rules, the space of third-order strategies contains  $2^{12} = 4,096$  elements.

Herein, we have only considered the dynamics for a small subset of third-order strategies (but the methods that we introduce in the following sections equally apply to all other strategies). The subset we have considered consists of the so-called leading-eight strategies (4). Under the assumption that all relevant information is shared publicly and that all players agree on everyone's reputation, past research has shown that if the whole population adopts any of the leading-eight strategies, the population will be fully cooperative and no other strategy can invade (5). However, the assumption of publicly shared



information seems rather stringent; it implies that the reputation repositories of all players agree at all times,  $\mathbf{m}_i(t) = \mathbf{m}_j(t)$  for all  $i, j$  and  $t$ . Instead, we explore the robustness of the leading-eight strategies when information is incomplete, noisy, and private. In that case, different players may hold different beliefs about their co-players, such that for any pair of players  $i$  and  $j$  there may be times for which  $\mathbf{m}_i(t) \neq \mathbf{m}_j(t)$ . To probe the robustness of the leading-eight strategies, we let each of these strategies compete with the two unconditional strategies ALLC and ALLD. Herein, we define these two strategies as ALLC =  $(\mathbf{1}; \mathbf{1})$  and ALLD =  $(\mathbf{0}; \mathbf{0})$ . We note that alternative strategy representations are possible: For example, we could define ALLC as the strategy that deems everyone as bad but cooperates anyway. Using such an alternative definition would not alter our qualitative results.

## 1.2 A formal description of the reputation dynamics

If the strategies  $(\alpha^i; \beta^i)$  are given for all players  $1 \leq i \leq N$ , we can describe the reputation dynamics of the population as a Markov chain. As the state of the Markov chain at time  $t$ , we collect all individual reputation repositories into an  $N \times N$  matrix

$$M(t) = \left( m_{ij}(t) \right). \quad [4]$$

We call  $M(t)$  the image matrix of the population at time  $t$ . Given the players' strategies and the current image matrix  $M = M(t)$  we can in principle calculate the transition probability  $h_{M, M'}$  that describes how likely we are to find the population in state  $M'$  at time  $t + 1$ . Because in every single time-step at most one column of the image matrix can change (the column with respect to the player who was randomly chosen to be the donor), we observe that  $h_{M, M'} = 0$  if  $M$  and  $M'$  differ in more than one column. Otherwise, when the image matrices  $M$  and  $M'$  differ in at most one column, the exact value of  $h_{M, M'}$  depends on the players' strategies, on the observation probability  $q$ , and on the error probability  $\varepsilon$ .

Over time, the average probability to observe a given image matrix  $M$  approaches an invariant distribution  $\mathbf{v} = (v_M)$  of the Markov chain  $H = (h_{M, M'})$ , such that  $\mathbf{v}H = \mathbf{v}$  and  $\sum_M v_M = 1$ . The entries  $v_M$  give the expected frequency with which we observe the image matrix  $M$  over the course of the indirect reciprocity game. Given the invariant distribution  $\mathbf{v}$ , we can calculate the average probability that player  $i$  considers  $j$  as good as

$$\bar{m}_{ij} = \sum_M v_M \cdot m_{ij}. \quad [5]$$

Moreover, we can calculate the average probability  $\bar{x}_{ij}$  with which player  $i$  cooperates with player  $j$  over the course of the game as

$$\bar{x}_{ij} = \sum_M v_M \beta_{m_{ii}, m_{ij}}^i. \quad [6]$$

Using this average cooperation probability, we define player  $i$ 's expected payoff  $\bar{\pi}_i$  as

$$\bar{\pi}_i = \frac{1}{N-1} \sum_{j \neq i} b \bar{x}_{ji} - c \bar{x}_{ij}. \quad [7]$$

Unfortunately, we note that the above approach quickly becomes computationally infeasible as the population becomes large. For a population of size  $N$ , the image matrix has  $N^2$  entries, implying that there are  $2^{N^2}$  possible image matrices.

Instead of calculating numerically exact payoffs according to Eq. [7], we have thus simulated the reputation dynamics to obtain estimates  $\hat{x}_{ij}$  for the average probability that player  $i$  cooperates with  $j$ . These estimates are then plugged into Eq. [7] to obtain estimated payoffs

$$\hat{\pi}_i = \frac{1}{N-1} \sum_{j \neq i} b\hat{x}_{ji} - c\hat{x}_{ij}. \quad [8]$$

By the theory of Markov chains, the estimated payoffs according to [8] converge to the true values [7] if the simulation is iterated for sufficiently many time steps. Unless noted otherwise, we have simulated the Markov chain for  $2 \cdot 10^6$  time steps, using a population of size  $N = 50$ . The respective simulations have been run with MATLAB; all scripts are provided in the **Appendix**.

Depending on the players' strategies, the Markov chain of the reputation dynamics does not need to be ergodic. In those cases, the players' average payoffs according to Eq. [8] may depend on their initial reputations. All our figures are based on the assumption that all players begin with a good reputation. To explore the robustness of our results, we have re-run the simulations in **Fig. 2** for two alternative scenarios. First we have run ten independent simulations in which the initial reputations are assigned randomly. For each of these ten simulations, we have obtained the same result as reported in the main text. Second, we have run a simulation assuming that all players start with a bad reputation. Again, we obtain the same result as in **Fig. 2** for seven out of the eight cases. Only for L7 the result is different. According to the assessment rule of L7 (**Tab. 1**), a bad donor can only gain a good reputation by cooperating against a good recipient. However, since there are no good recipients to start with, it is impossible for donors to gain a good standing in the eyes of an L7 observer. Hence, all L7 players keep their initial bad assessment of all co-players. We conclude that under such an initial assignment of reputations, L7 does not yield any cooperation at all.

### 1.3 Recovery analysis after single disagreements

While we use simulations to explore the general reputation dynamics under noisy private information, we can derive analytical results in the limiting case that the players' actions are perfectly observable ( $q = 1$ ) and that perception errors are rare ( $\varepsilon \rightarrow 0$ ). To this end, we consider a homogeneous population of players who all apply the same leading-eight strategy (as given in **Tab. 1**). We assume that initially, all players perceive everyone as good; only player 1 perceives player 2 as bad, possibly because of an error (**Fig. S3a**). That is, the initial image matrix is given by  $M(0) = M^0$  with entries

$$m_{ij}^0 = \begin{cases} 0 & \text{if } i = 1, j = 2 \\ 1 & \text{otherwise.} \end{cases} \quad [9]$$

We assume that in subsequent rounds, no further errors occur and that all individuals observe all co-players' interactions. We say that the population recovers from a single disagreement, if it returns to the state where all players have a good reputation starting from the state  $M^0$ . In the following we are interested in the following two quantities, depending on the applied leading-eight strategy  $L_i$ :

1. The population's recovery probability  $\rho_i$
2. The expected time until recovery  $\tau_i$  (conditioned on that the population actually recovers)

The following Proposition simplifies the corresponding analysis.

**Proposition 1.** *Consider the indirect reciprocity game for a population in which everyone applies the same leading-eight strategy  $L_i$ . Moreover, assume that the initial image matrix is  $M(0) = M^0$  as defined in [9], and let  $M(t)$  denote the image matrix at some subsequent time  $t > 0$  according to the process with perfect observation and no noise,  $q = 1$  and  $\varepsilon = 0$ . Then  $M(t) \in \mathcal{M}$ , where  $\mathcal{M}$  is the set of all image matrices that satisfy the following three conditions*

$$(i) \ m_{ii} = 1 \text{ for all } i, \quad (ii) \ m_{ij} = 1 \text{ for all } i, j \geq 2, \quad (iii) \ m_{i1} = m_{j1} \text{ for all } i, j \geq 2. \quad [10]$$

All proofs are provided in the **Appendix**. The above Proposition guarantees that in the process with perfect observation and no noise, (i) all players think of themselves as good, (ii) the players  $2 \leq i, j \leq N$  consider each other as good, and (iii) all players  $2 \leq i, j \leq N$  have the same opinion about player 1.

Proposition 1 is useful because it allows us to consider a simplified state space. Instead of considering the space of all image matrices  $M$ , in the following we consider the space of all tuples  $(r, k)$  with  $r \in \{0, 1\}$  and  $k \in \{0, \dots, N-1\}$ . The value of  $r$  refers to the reputation of player 1 from the perspective of all other players (due to Proposition 1 (iii) all other players agree on player 1's reputation). We use  $r = 1$  to indicate that player 1 is perceived as good, and  $r = 0$  to indicate that she is perceived as bad. The value of  $k$  refers to the number of co-players that player 1 considers as good (due to Proposition 1, we can treat all other players as equivalent). In this reduced state space, the initial state thus corresponds to the pair  $(1, N-2)$ , and the full recovery state is  $(1, N-1)$ , see **Fig. S3a** for an illustration.

Let  $f^i(r, k; r', k')$  denote the transition probabilities for the reduced state space; the value of  $f^i(r, k; r', k')$  corresponds to the probability that a population of  $L_i$  players moves from state  $(r, k)$  to  $(r', k')$  in one round. We can deduce these probabilities as follows:

Transition  $(1, k) \rightarrow (1, k+1)$ . This case can only occur if a player  $i > 1$  is chosen to be the donor who is perceived as bad by player 1. Given that the current state is  $(1, k)$ , it follows from Proposition 1 that the donor considers everyone as good, and hence she cooperates. If player 1 considers the receiver to be good, this leads her to assign a good reputation to the donor, independent of the applied leading-eight strategy  $L_i$ . Otherwise, if player 1 considers the receiver to be bad, the donor only obtains a good reputation for  $L_1, L_2, L_3$ , and  $L_5$ . Therefore, the corresponding transition

probability is

$$f^i(1, k; 1, k+1) = \begin{cases} \frac{N-k-1}{N} & \text{if } i \in \{1, 2, 3, 5\} \\ \frac{N-k-1}{N} \frac{k+1}{N-1} & \text{if } i \in \{4, 6, 7, 8\}. \end{cases} \quad [11]$$

Transition  $(1, k) \rightarrow (1, k-1)$ . This case can only occur if a player  $i > 1$  is randomly chosen to act as the donor who is perceived as good by player 1. Similar to before, player  $i$  will always cooperate, which is only considered as bad by player 1 if the receiver is considered as bad by player 1 and if the applied strategy is either  $L_2, L_5, L_6$ , or  $L_8$ . Therefore, the transition probability is

$$f^i(1, k; 1, k-1) = \begin{cases} 0 & \text{if } i \in \{1, 3, 4, 7\} \\ \frac{k}{N} \frac{N-k-1}{N-1} & \text{if } i \in \{2, 5, 6, 8\}. \end{cases} \quad [12]$$

Transition  $(1, k) \rightarrow (0, k)$ . This case can only occur if player 1 is chosen to be the donor, and if player 1 defects against the receiver (which in turn requires player 1 to consider the receiver as bad). The corresponding transition probability is

$$f^i(1, k; 0, k) = \frac{1}{N} \frac{N-k-1}{N-1}. \quad [13]$$

Transition  $(0, k) \rightarrow (0, k+1)$ . This case requires that a player  $i > 1$  is chosen to be the donor who is considered as bad by player 1. This donor cooperates, unless the randomly chosen receiver happens to be player 1 (who is bad from the perspective of all other players). Thus, player 1 considers the donor as good after this round unless the receiver is player 1, or the receiver is a group member that is considered as bad by player 1 and the applied leading-eight strategy is  $L_4, L_6, L_7$ , or  $L_8$ . Hence, we obtain

$$f^i(0, k; 0, k+1) = \begin{cases} \frac{N-k-1}{N} \frac{N-2}{N-1} & \text{if } i \in \{1, 2, 3, 5\} \\ \frac{N-k-1}{N} \frac{k}{N-1} & \text{if } i \in \{4, 6, 7, 8\}. \end{cases} \quad [14]$$

Transition  $(0, k) \rightarrow (0, k-1)$ . This case requires that a player  $i > 1$  is chosen to be the donor who player 1 considers as good. To become bad in player 1's eyes, this donor then either needs to defect against player 1, or he needs to cooperate against a receiver who is considered as bad by player 1 (provided that the applied leading-eight strategy is  $L_2, L_5, L_6$ , or  $L_8$ ). The transition probability becomes

$$f^i(0, k; 0, k-1) = \begin{cases} \frac{k}{N} \frac{1}{N-1} & \text{if } i \in \{1, 3, 4, 7\} \\ \frac{k}{N} \frac{N-k}{N-1} & \text{if } i \in \{2, 5, 6, 8\}. \end{cases} \quad [15]$$

Transition  $(0, k) \rightarrow (1, k)$ . This case requires player 1 to be the donor, and that player 1 cooperates with

her co-player. The corresponding probability is

$$f^i(0, k; 1, k) = \frac{1}{N} \frac{k}{N-1}. \quad [16]$$

All other transitions from  $(r, k)$  to  $(r', k')$  have transition probability  $f^i(r, k; r', k') = 0$ . We note that for this reduced Markov chain, the recovery state  $(1, N-1)$  is absorbing. Once this state is reached, it cannot be left anymore, because  $f^i(1, N-1; 1, N-2) = 0$  and  $f^i(1, N-1; 0, N-1) = 0$  for all  $i$ . However, for populations that apply one of the four leading-eight strategies  $L_4, L_6, L_7, L_8$ , there exists another absorbing state, which is  $(0, 0)$ . This state corresponds to a full segregation state: player 1 considers everyone else as bad, whereas all other players consider player 1 to be bad (**Fig. S3a**). Whether the population is able to recover from a single disagreement thus depends on how likely the full segregation state is reached. We obtain the following three cases.

**Proposition 2** (Recovery probabilities).

Suppose the population applies the leading-eight strategy  $L_i$ .

1. For  $i \in \{1, 2, 3, 5\}$ , the recovery probability is  $\rho_i = 1$ .
2. For  $i \in \{4, 7\}$ , the recovery probability satisfies  $1 - 2/(N-1)! \leq \rho_i < 1$ .
3. For  $i \in \{6, 8\}$ , the recovery probability is  $\rho_i = 1 - 1/N$ .

With respect to the time it takes the population to recover from a single disagreement, we obtain a similar case distinction.

**Proposition 3** (Expected time until recovery).

Suppose the population applies the leading-eight strategy  $L_i$ .

1. For  $i \in \{1, 3, 4, 7\}$ , the expected time until recovery  $\tau_i$  is of order  $\Theta(N)$ .
2. For  $i \in \{2, 5\}$ , the expected time until recovery  $\tau_i$  is of order  $\Theta(N \log N)$ .
3. For  $i \in \{6, 8\}$ , the expected time until recovery is  $\tau_i = N \cdot H_N - N$ , where  $H_N = \sum_{n=1}^N \frac{1}{n}$  is the  $N$ -th harmonic number.

While some of the results in Proposition 3 only address the asymptotic behavior as the population becomes large, we can numerically compute exact solutions for the expected recovery time for all population sizes, based on the transition probabilities  $f^i(r, k; r', k')$  specified above. The respective solutions are shown in **Fig. S3b**.

The above results suggest that with respect to recovery from single disagreements, there are four groups of leading-eight strategies. The strategies  $L_1$  and  $L_3$  are most robust: after a single disagreement, they are guaranteed to recover, and the expected recovery time is linear in the population size. These two strategies have in common that a cooperating donor is always perceived as good (even if the donor or the receiver had a bad reputation to start with). Due to this property, a single individual who is perceived as

bad can immediately regain a good reputation once it is chosen to act as a donor (which takes an expected time of  $N$  rounds).

A similar observation holds for the two strategies  $L_4$  and  $L_7$ : a single individual with bad reputation can always regain a good reputation by cooperating, independent of the identity of the receiver. Unlike the first two strategies, however,  $L_4$  and  $L_7$  can end up in the full segregation state. However, this requires a rather specific sequence of events: before player 2 can regain his good reputation in the eyes of player 1, it needs to be the case that player 1 is chosen to act as the donor with respect to player 2 as the recipient. In such a scenario, player 1 would defect, yielding her a bad reputation from the viewpoint of all other players. To end up in the full segregation state, all other players would then have to be chosen to act as a donor against player 1. As the population grows bigger, this sequence of events becomes increasingly unlikely, and the recovery probability quickly approaches 1.

Recovery is guaranteed for the two strategies  $L_2$  and  $L_5$ , but in contrast to the first four strategies, recovery may take substantially longer. The longer recovery time is due to the property of these strategies that cooperation against an individual perceived as bad is itself perceived as bad. Thus, if player 2 is randomly chosen to be the recipient of the next indirect reciprocity interaction, the donor is guaranteed to obtain a bad reputation in the eyes of someone (either because the donor is an individual  $i \geq 2$ , who is then perceived as bad by player 1; or because the donor is player 1, who is then perceived as bad by everyone else).

The above recovery analysis confirms that  $L_6$  and  $L_8$  perform worst in the presence of single disagreements. Recovery is not guaranteed, and even if it occurs, it may take a substantial time until recovery. Both aspects are due to the relative ease with which these two strategies assign a bad reputation to other population members: a donor who cooperates against a receiver who is perceived as bad is guaranteed to end up with a bad reputation (independent of whether the donor was perceived as good or bad before). Under these conditions, bad reputations can proliferate comparably quickly.

## 2 Evolutionary dynamics

### 2.1 Description of the evolutionary process

While the previous section has assumed that the players' strategies are fixed, we now describe a simple process that allows us to explore which dynamics arises once players may change their strategies over time. To allow for a transparent treatment, we assume that strategy adaptation occurs on a time scale that is slow compared to the reputation dynamics.

Specifically, we assume that the strategies in the population change according to a simple imitation process (6). For this process, we again consider a population of  $N$  individuals where each individual  $i$  may have its own strategy  $(\alpha^i; \beta^i)$ . In each time step of the evolutionary process, one individual  $i$  is chosen at random to update her strategy. There are two ways to do so: with probability  $\mu$  (corresponding to the mutation rate), she adopts a random strategy (with all remaining strategies having the same probability to be picked). With the remaining probability  $1 - \mu$ , she seeks for a role model instead, by randomly choosing another individual  $j$  from the population. Suppose player  $i$ 's payoff is given by  $\hat{\pi}_i$

and  $j$ 's payoff is  $\hat{\pi}_j$ , where  $\hat{\pi}_i$  and  $\hat{\pi}_j$  are given by the estimated payoffs according to Eq. [8]. We assume that the probability that  $i$  adopts  $j$ 's strategy is given by the Fermi function (7),

$$P(\hat{\pi}_j, \hat{\pi}_i) = \frac{1}{1 + \exp(-s(\hat{\pi}_j - \hat{\pi}_i))}. \quad [17]$$

The parameter  $s \geq 0$  is called the strength of selection. It determines how much players value payoffs when adopting new strategies. For  $s = 0$  we obtain  $P(\hat{\pi}_j, \hat{\pi}_i) = 1/2$ , and imitation occurs essentially at random. For  $s > 0$ , the imitation probability  $P(\hat{\pi}_j, \hat{\pi}_i)$  is monotonically increasing in the payoff difference  $\hat{\pi}_j - \hat{\pi}_i$ . That is, the more there is to gain for player  $i$ , the more likely she is to abandon her old strategy and to imitate player  $j$ 's strategy instead.

This elementary updating process, involving mutation and imitation, is then iterated over many time steps. As a result, we obtain an ergodic process on the space of all possible population compositions. This process has again a unique invariant distribution, to which we refer as the selection-mutation equilibrium. The abundance of each strategy in this equilibrium, and the corresponding average cooperation rate, can always be obtained by simulating the above process over a sufficiently long time span. However, in certain special cases, the selection-mutation equilibrium can be computed more efficiently. These more efficient algorithms either require a small or intermediate-sized population in which only a few different strategies are available, or that mutations are rare. In the following two sections we describe these algorithms in more detail.

Here, we have phrased the evolutionary process as an imitation dynamics. However, similar results can be obtained if one assumes that strategies are genetically encoded and inherited in a birth-death process (8), where the fitness of an individual is given by an exponential function of its payoffs,  $\exp(s\hat{\pi}_i)$ .

## 2.2 Selection-mutation equilibrium in populations with few strategies

Throughout the main text, we have focused on a population of intermediate size  $N$  that has access to at most  $k = 3$  different strategies. The possible states of the population are elements of the set

$$\Delta_N^k := \left\{ \mathbf{n} = (n_1, \dots, n_k) \in \mathbb{N}^k \mid \sum_{i=1}^k n_i = N \right\}. \quad [18]$$

The entries  $n_i$  of each vector represent how many players currently apply strategy  $i$ . The number of such population compositions is

$$|\Delta_N^k| = \binom{N+k}{k}. \quad [19]$$

The evolutionary process described in Section 2.1 defines a Markov chain with state space  $\Delta_N^k$ . For two population states  $\mathbf{n}, \mathbf{n}' \in \Delta_N^k$ , the transition probability to move from  $\mathbf{n}$  to  $\mathbf{n}'$  in one step of the process

is given by

$$w_{\mathbf{n},\mathbf{n}'} = \begin{cases} \frac{n_i}{N} \left( \frac{\mu}{k-1} + (1-\mu) \frac{n_j}{N} P(\hat{\pi}_j, \hat{\pi}_i) \right) & \text{if } n'_i = n_i - 1, n'_j = n_j + 1, n'_l = n_l \text{ for } l \notin \{i, j\} \\ 1 - \sum_{j \neq i} \frac{n_i}{N} \left( \frac{\mu}{k-1} + (1-\mu) \frac{n_j}{N} P(\hat{\pi}_j, \hat{\pi}_i) \right) & \text{if } \mathbf{n} = \mathbf{n}' \\ 0 & \text{otherwise.} \end{cases} \quad [20]$$

Provided that neither the population size  $N$  nor the number of strategies  $k$  is prohibitively large, the transition matrix  $W = (w_{\mathbf{n},\mathbf{n}'})$  can be computed explicitly. By computing the normalized left eigenvector of  $W$  (with respect to eigenvalue 1), we obtain the selection-mutation equilibrium over an evolutionary timescale. We have implemented this algorithm using MATLAB; the corresponding code is provided in the **Appendix**.

### 2.3 Selection-mutation equilibrium in the limit of rare mutations

Alternatively, in arbitrarily large populations with an arbitrary number of available strategies, we can still calculate exact strategy abundances in the selection-mutation equilibrium if mutations are sufficiently rare (9–11). In that case, the population will find itself in a homogeneous state most of the time, in which all individuals adopt the same strategy  $i$ . Only occasionally, a mutant strategy  $j$  arises. This mutant then either reaches fixation in the population, or it goes extinct. The mutant's fixation probability  $p_{ij}$  can be calculated explicitly (12, 13),

$$p_{ij} = \frac{1}{1 + \sum_{l'=1}^{N-1} \prod_{l=1}^{l'} \exp(-s(\hat{\pi}_j(l) - \hat{\pi}_i(l)))}. \quad [21]$$

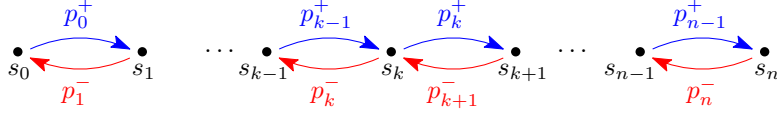
Here,  $\hat{\pi}_j(l)$  and  $\hat{\pi}_i(l)$  refer to the payoffs of a mutant and a resident, provided that the number of mutants in the population is  $l$ . Using Eq. [21], the selection-mutation equilibrium of the evolutionary process can be computed by considering a reduced Markov chain with  $k$  states (corresponding to the homogeneous populations in which every player applies the same strategy). The probability to move from state  $i$  to state  $j$  is given by  $p_{ij}/(k-1)$ . The invariant distribution of this reduced Markov chain approximates the invariant distribution of the evolutionary process as  $\mu$  becomes small (9). Again, we have implemented this algorithm in MATLAB; the code is given in the **Appendix**.

## Appendix

### Proofs of the recovery analysis

*Proof of Proposition 1.* For the proof, we consider the Markov chain  $H = (h_{M,M'})$ , as defined in Section 1.2, for the limiting case  $\varepsilon=0$  and  $q=1$ . We show that the set  $\mathcal{M}$  of image matrices that satisfy the three properties in [10] is invariant. That is, let  $M \in \mathcal{M}$  be arbitrary and suppose  $h_{M,M'} > 0$  for some image matrix  $M'$ . Then also  $M' = (m'_{ij})$  satisfies the three characteristic properties [10].





**Figure A1: One-dimensional random walk.** A discrete-time Markov chain with  $n+1$  states labelled  $s_0, \dots, s_n$  arranged in a line and with transition probabilities  $p_k^+ : s_k \rightarrow s_{k+1}$  and  $p_k^- : s_k \rightarrow s_{k-1}$ .

(i)  $m'_{ii} = 1$  for all  $i$ . Since  $M \in \mathcal{M}$ , initially all players consider themselves as good. All leading-eight strategies have the property that the strategy's action rule prescribes an action that lets a good donor maintain her good reputation in her own eyes, independent of which reputation she assigns to the recipient. Thus all players keep considering themselves as good after one interaction; either they do not need to make a decision (because they were not chosen to act as the donor), or they choose an action they themselves evaluate as good.

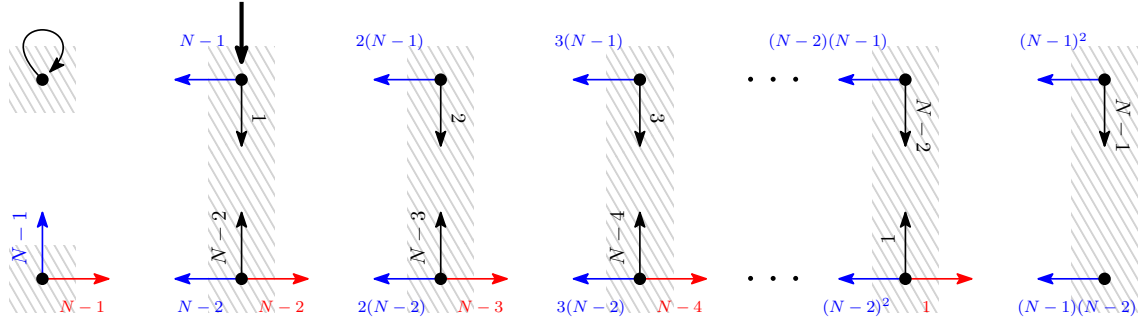
(ii) and (iii)  $m'_{ij} = 1$  and  $m'_{i1} = m'_{j1}$  for all  $i, j \geq 2$ . Since  $M \in \mathcal{M}$ , all players  $i, j \geq 2$  initially agree on the reputations of all population members. Because they all apply the same assessment rule and observation errors are excluded, they also agree on how the donor's action in the subsequent interaction needs to be assessed. This shows  $m'_{il} = m'_{jl}$  for all  $i, j \geq 2$  and all  $l$ . Moreover, since all players  $i, j \geq 2$  consider each other as good initially, and since their common action rule only lets them choose actions that let them keep their good reputation, we conclude  $m'_{ij} = 1$  for  $i, j \geq 2$ .

□

*Proof of Propositions 2 and 3.* In the following, we provide the proofs of Propositions 2 and 3 by considering each of the four different cases  $\{L1, L3\}$ ,  $\{L2, L5\}$ ,  $\{L4, L7\}$ ,  $\{L6, L8\}$  individually. For a given leading-eight strategy  $L_i$ , we consider the respective Markov chain  $M_i$  with  $2N$  states  $s_{r,k}$ . The value of  $r \in \{0, 1\}$  corresponds to the reputation of player 1 from the perspective of all other players. The value of  $k \in \{0, 1, \dots, N-1\}$  corresponds to the number of co-players that player 1 considers as good. Then the recovery probability  $\rho_i$  is the probability that a random realization of the process (a *random trace*) starting in a state  $s_{1,N-2}$  reaches the absorbing state  $s_{1,N-1}$ . The expected recovery time  $\tau_i$  is the expected number of steps until a random trace starting in a state  $s_{1,N-2}$  reaches the state  $s_{1,N-1}$ , conditioned on it doing so. The main idea of the proof is to make use of a coupling argument. That is, to each two-dimensional random walk defined by the Markov chain  $M_i$ , we associate a simpler, one-dimensional random walk. For the one-dimensional random walk, the absorption probabilities and conditional absorption times can be calculated explicitly, and they can serve as upper (or lower) bounds for the respective quantities in the two-dimensional random walk. To this end, let us recall the following result.

**Proposition 4** (One-dimensional random walk, see Chapter 7.7, proof of Theorem 7.1 in (14)).

Let  $M$  be a discrete-time Markov chain with  $n+1$  states  $s_0, \dots, s_n$  and transition probabilities  $p_i^+ : s_i \rightarrow s_{i+1}$  ( $i = 0, \dots, n-1$ ) and  $p_i^- : s_i \rightarrow s_{i-1}$  ( $i = 1, \dots, n$ ), as illustrated in Fig. A1.



**Figure A2:** Markov chain  $M_1 = M_3$ . The transition probabilities are normalized by  $N(N-1)$ . Self-loops in non-absorbing states are not shown.

1. If both states  $s_0$  and  $s_n$  are absorbing (i.e.  $p_0^+ = p_n^- = 0$ ) then the probability  $\rho(k)$  of reaching  $s_n$  before reaching  $s_0$  when starting at  $s_k$  is given by

$$\rho(k) = \frac{1 + r_1 + r_1 r_2 + \cdots + r_1 \cdots r_{k-1}}{1 + r_1 + r_1 r_2 + \cdots + r_1 \cdots r_{n-1}}, \quad [22]$$

where  $r_i = p_i^- / p_i^+$  for each  $i = 1, \dots, n-1$ .

2. If there is no absorbing state except, possibly,  $s_0$  then the expected number of time steps  $t_{k,k-1}$  to reach state  $s_{k-1}$  from state  $s_k$  is given by

$$t_{k,k-1} = \frac{1}{p_k^-} + \frac{p_k^+}{p_{k+1}^- p_k^-} + \cdots + \frac{p_k^+ \cdots p_{n-1}^+}{p_n^- \cdots p_k^-} = \frac{1}{p_k^-} \left( 1 + \sum_{i=1}^{n-k} \prod_{j=1}^i \frac{p_{k+j-1}^+}{p_{k+j}^-} \right). \quad [23]$$

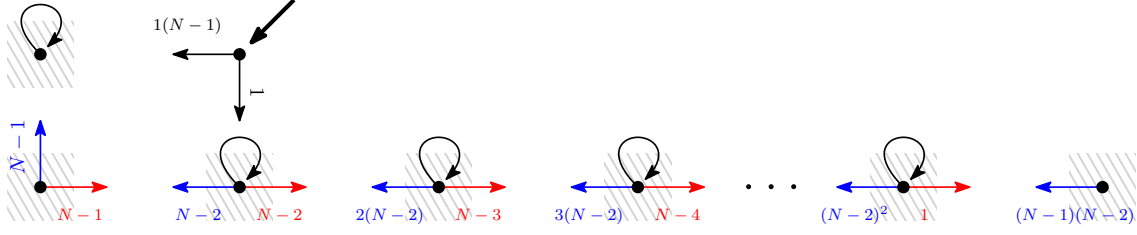
Moreover, the expected number of time steps  $t_{k,l}$  to reach state  $s_l$  from state  $s_k$ , with  $l < k$ , is

$$t_{k,l} = t_{k,k-1} + t_{k-1,k-2} + \cdots + t_{l+1,l} = \sum_{i=l+1}^k t_{i,i-1}. \quad [24]$$

We use standard notation  $o(\cdot)$  and  $\Theta(\cdot)$  for strict asymptotic upper bound and for asymptotically tight bound, ignoring the constant factors. Hence, for example, we have  $1/n = o(1)$  and  $2n+1 = \Theta(n)$ , because for large  $n$  we have  $1/n \ll 1$  whereas  $\frac{2n+1}{n}$  tends to a constant. See Section 1.3 of (15) for a detailed treatment. For asymptotic results, we assume  $N \geq 4$ .

Recovery analysis for  $L_1 = L_3$ . The Markov chain  $M_1 = M_3$  is depicted in Fig. A2.

- **Recovery probability:**  $M_1$  has a unique absorbing state  $s_{1,N-1}$ . Moreover, starting from any other state, there is a positive probability to reach the state  $s_{1,N-1}$ . Hence, it follows that  $\rho_1 = 1$ .
- **Recovery time:** To get a lower bound on the recovery time, we note that reaching  $s_{1,N-1}$  requires at least one non-self-looping transition. Starting from the state  $s_{1,N-2}$ , the expected time until the first



**Figure A3:** Markov chain  $M'_1$ . The transition probabilities are normalized by  $N(N-1)$ . By coupling,  $\tau'_1 \geq \tau_1$ .

non-self-looping transition occurs is  $\frac{N(N-1)}{N-1+1} = N-1$ . Hence  $\tau_1 \geq N-1$ .

In the following, we show  $\tau_1 \leq N+7$ , which establishes the desired  $\tau_1 = \Theta(N)$ . To this end, consider the Markov chain  $M'_1$  obtained from  $M_1$  by erasing states  $s_{1,k}$  for  $k \leq N-2$ , and by replacing transitions  $s_{0,k} \rightarrow s_{1,k}$  with self-loops (see Fig. A3).

First, we argue that  $\tau_1 \leq \tau'_1$ , where  $\tau'_1$  is the expected number of steps to reach  $s_{1,N-1}$  from  $s_{1,N-2}$  in  $M'_1$ . To this end, consider an arbitrary trace  $T$  in  $M_1$ . If  $T$  never takes any transition  $s_{0,k} \rightarrow s_{1,k}$ , for  $k < N-1$ , then we can associate the identical trace  $T'$  in  $M'_1$  to  $T$ . Otherwise, if there is a moment when  $T$  moves from  $s_{0,k}$  to  $s_{1,k}$  for some  $k < N-1$ , the associated trace  $T'$  has a self-loop at  $s_{0,k}$ . Since for each  $k < N-1$  the transition  $s_{1,k} \rightarrow s_{1,k+1}$  has higher probability than the transition  $s_{0,k} \rightarrow s_{0,k+1}$ , we can then couple traces  $T$  and  $T'$  such that, from that point on,  $T$  is not to the right of or below  $T'$  (formally, at any time after that moment, the  $k$ -coordinate of  $T$  is larger or equal to the  $k$ -coordinate of  $T'$ , and the  $r$ -coordinate of  $T$  is larger or equal to the  $r$ -coordinate of  $T'$ ). In particular, if  $T'$  has reached  $s_{1,N-1}$  then so did  $T$ , and the inequality  $\tau_1 \leq \tau'_1$  follows.

It remains to prove that  $\tau'_1 \leq N+7$ . We use Proposition 4. Let  $x$  be the expected number of steps to reach  $s_{1,N-1}$  from  $s_{0,N-2}$  in  $M'_1$ . Then  $\tau'_1 = (N-1) + \frac{N-1}{N} \cdot 0 + \frac{1}{N} \cdot x$ . By Eq. [24],  $x$  takes the form  $x = t_{2,1} + t_{1,0}$ . The quantities  $t_{2,1}$  and  $t_{1,0}$  are calculated using Eq. [23]. For the first quantity, we get

$$\begin{aligned} t_{2,1} &= N(N-1) \left( \frac{1}{N-2} + \frac{N-2}{2!(N-2)^2} + \frac{(N-2)(N-3)}{3!(N-2)^3} + \cdots + \frac{(N-2)!}{(N-1)!(N-2)^{N-1}} \right) \\ &\leq \frac{N(N-1)}{N-2} \left( 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(N-1)!} \right) \leq \frac{N(N-1)}{N-2} \cdot (e-1), \end{aligned}$$

where we have made use of the definition of Euler's number,  $e := \sum_{i=0}^{\infty} 1/i!$ . For the second quantity, we similarly obtain

$$\begin{aligned} t_{1,0} &= N(N-1) \left( \frac{1}{N-1} + \frac{N-1}{(N-1)(N-2)} + \frac{(N-1)(N-2)}{(N-1)2!(N-2)^2} + \cdots + \frac{(N-2)!}{(N-1)(N-1)!(N-2)^{N-1}} \right) \\ &\leq \frac{N(N-1)}{N-2} \left( 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{(N-1)!} \right) \leq \frac{N(N-1)}{N-2} \cdot e \end{aligned}$$

Hence  $x = t_{2,1} + t_{1,0} \leq \frac{N(N-1)}{N-2} \cdot (2e-1)$  and, for  $N \geq 4$ ,

$$\tau_1' = N - 1 + \frac{1}{N} \cdot x \leq N - 1 + \frac{N-1}{N-2} \cdot (2e-1) = N + 2(e-1) + \frac{2e-1}{N-2} < N + 7,$$

as desired.

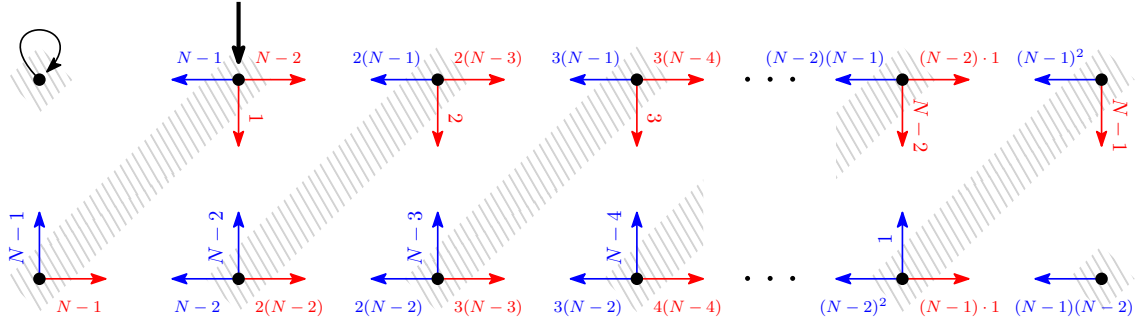
Recovery analysis for  $L_2 = L_5$ . The Markov chain  $M_2 = M_5$  is depicted in Fig. A4.

- **Recovery probability:** As before, since there is a unique absorbing state that can be reached from any initial state,  $\rho_2 = 1$ .
- **Recovery time:** To show that the expected number of steps  $\tau_2$  is of the order  $\Theta(N \cdot \log N)$ , we consider two simpler Markov chains  $M_2^+$  (resp.  $M_2^-$ ) with expected number of steps  $\tau_2^+$  (resp.  $\tau_2^-$ ). We then prove that  $\tau_2^+ \geq \tau_2 \geq \tau_2^-$  and that both  $\tau_2^+$  and  $\tau_2^-$  are of the order  $\Theta(N \cdot \log N)$ .

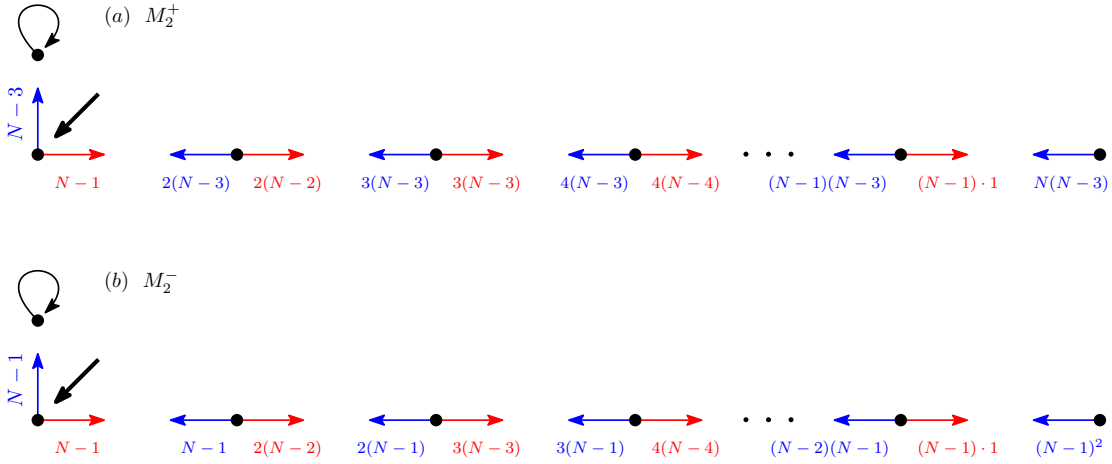
Intuitively,  $M_2^+$  is obtained by identifying states  $s_{1,k}$  and  $s_{0,k+1}$  for each  $k < N-1$  (that is, we identify all states that have the same number of individuals that are universally considered as good). For the transitions of  $M_2^+$ , we take the transition probabilities as in Fig. A5(a). In doing so, we decrease the total probability to move either to the left or upwards in each state, while leaving the total probability to move to the right or downwards unchanged. As a consequence, any trace  $T^+$  in  $M_2^+$  can be coupled with a trace  $T$  in  $M_2$  such that if  $T^+$  is in a state  $s_{r^+,k^+}$  at time  $t$ , and  $T$  is in a state  $s_{r,k}$ , then  $r^+ + k^+ \leq r + k$ . Hence, if  $T^+$  has reached  $s_{1,N-1}$  then so did  $T$  and the inequality  $\tau_2^+ \geq \tau_2$  follows. For  $M_2^-$  we proceed analogously: we use the same identification of states, but this time the probability to move left in Fig. A5(b) is larger or equal to the corresponding probability to move left or up in Fig. A4. As a consequence,  $\tau_2 \geq \tau_2^-$ .

Finally, we compute  $\tau_2^+, \tau_2^-$ . By Eq. [23], we obtain

$$\begin{aligned} \tau_2^+ &= N(N-1) \left( \frac{1}{N-3} + \frac{N-1}{2!(N-3)^2} + \frac{1 \cdot 2 \cdot (N-1)(N-2)}{3!(N-3)^3} + \cdots + \frac{1 \cdot 2 \cdots (N-1) \cdot (N-1)(N-2) \cdots 1}{N!(N-3)^N} \right) \\ &= \frac{N(N-1)}{N-3} \left( 1 + \frac{1}{2} \cdot \frac{N-1}{N-3} + \frac{1}{3} \cdot \frac{(N-1)(N-2)}{(N-3)^2} + \cdots + \frac{1}{N} \cdot \frac{(N-1)!}{(N-3)^{N-1}} \right) \\ &\leq \frac{N(N-1)^2(N-2)}{(N-3)^3} \left( 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N} \right) = \Theta(N \log N), \end{aligned}$$



**Figure A4:** Markov chain  $M_2 = M_5$ . The transition probabilities are normalized by  $N(N-1)$ . Self-loops in non-absorbing states are not shown.



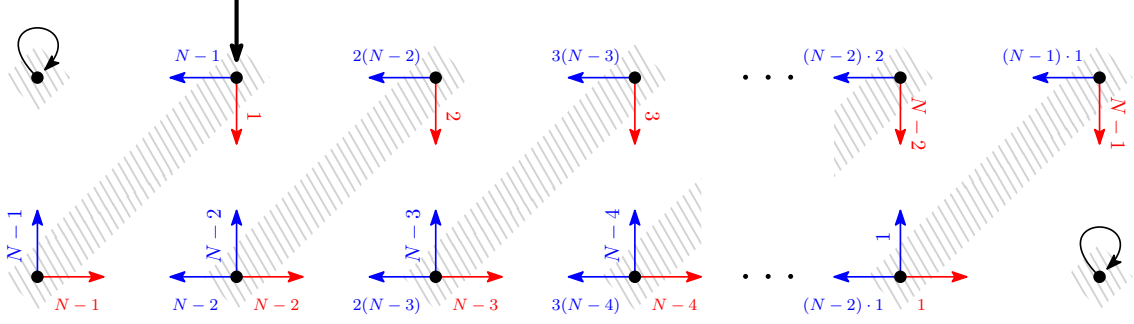
**Figure A5:** Markov chains  $M_2^+$  and  $M_2^-$ . The transition probabilities are normalized by  $N(N-1)$ . By coupling,  $\tau_2^+ \geq \tau_2 \geq \tau_2^-$ .

$$\begin{aligned}
\tau_2^- &= N(N-1) \left( \frac{1}{N-1} + \frac{N-1}{2!(N-1)^2} + \frac{1 \cdot 2 \cdot (N-1)(N-2)}{3!(N-1)^3} + \dots + \frac{1 \cdot 2 \cdot \dots \cdot (N-1) \cdot (N-1)(N-2) \cdot \dots \cdot 1}{N!(N-1)^N} \right) \\
&\geq N \left( 1 + \frac{1}{2} \cdot \frac{N-1}{N-1} + \frac{1}{3} \cdot \frac{(N-1)(N-2)}{(N-1)^2} + \dots + \frac{1}{\sqrt{N}} \cdot \frac{(N-1) \dots (N-\sqrt{N})}{(N-1)^{\sqrt{N}}} \right) \\
&\geq N \left( 1 + \frac{1}{2} \frac{N-1}{N-1} + \frac{1}{3} \frac{N-3}{N-1} + \dots + \frac{1}{\sqrt{N}} \frac{N-\frac{1}{2}N}{N-1} \right) \geq N \cdot \frac{1}{2} \cdot H_{\sqrt{N}} \geq \frac{1}{4} N \log N = \Theta(N \log N),
\end{aligned}$$

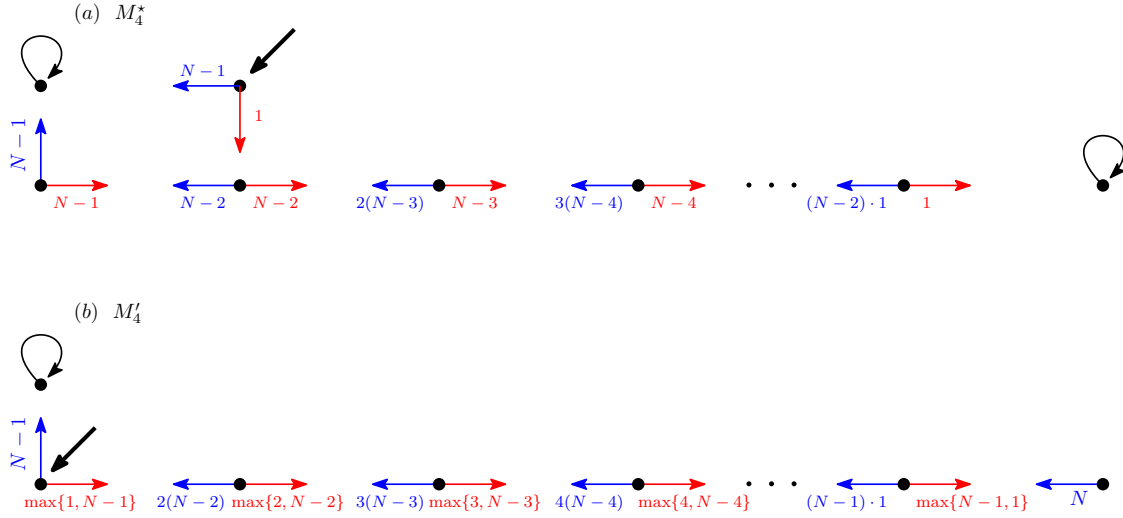
where we only took the sum of the first  $\sqrt{N}$  terms and used that  $1 + 2 + \dots + \sqrt{N} \approx \frac{1}{2}N$  and that  $\log \sqrt{N} = \frac{1}{2} \log(N)$ .

Recovery analysis for  $L_4 = L_7$ . The Markov chain  $M_4 = M_7$  is depicted in Fig. A6.

- **Recovery probability:**  $M_4$  has two absorbing states,  $s_{0,0}$  and  $s_{1,N-1}$ , and both states can be reached with positive probability from any other state. Nevertheless, we show in the following that  $\rho_4$  quickly



**Figure A6:** Markov chain  $M_4 = M_7$ . The transition probabilities are normalized by  $N(N-1)$ . Self-loops in non-absorbing states are not shown.



**Figure A7:** Markov chains  $M_4^*$  and  $M_4'$ . The transition probabilities are normalized by  $N(N-1)$ . By coupling,  $\tau_4 \leq \tau_4'$ .

approaches 1 as the population size increases, because  $\rho_4 \geq 1 - 2/(N-1)!$ .

To obtain this lower bound for  $\rho_4$ , we consider the Markov chain  $M_4^*$  that is obtained from  $M_4$  by erasing states  $s_{1,k}$  for  $k < N-1$ , and replacing the corresponding transitions of the form  $s_{0,k} \rightarrow s_{1,k}$  by self-loops. Formally, we define  $M_4^*$  to be as in Fig. A7(a). For the coupling, note that any time a trace  $T$  in  $M_4$  takes a transition of the form  $s_{0,k} \rightarrow s_{1,k}$ , the corresponding trace  $T^*$  in  $M_4^*$  waits until (if ever)  $T$  reaches state  $s_{r,k}$  with  $r = 0$  again. At that point,  $T$  is to the left of  $T^*$  or in the same state. Hence if  $T^*$  ever reaches  $s_{1,N-1}$  then so did  $T$  and we have  $\rho_4 \geq \rho_4'$ .

Denoting by  $x$  the probability of reaching  $s_{0,0}$  from  $s_{0,N-2}$  (before reaching  $s_{1,N-1}$ ), we have  $1 - \rho_4' = \frac{N-1}{N} \cdot 0 + \frac{1}{N} \cdot x = x/N$  and by Proposition 4 we compute

$$x = \frac{1 + 1}{1 + 1 + 1! + 2! + 3! + \dots + (N-2)!} \leq \frac{2}{(N-2)!}$$

implying that  $\rho_4 \geq \rho_4' = 1 - x/N \geq 1 - 2/(N-1)!$  as desired.

- **Recovery time:** since reaching  $s_{1,N-1}$  requires taking at least one non-self-looping transition, we immediately get  $\tau_4 \geq \frac{N(N-1)}{N-1+1} = N-1$ . In the following, we focus on the upper bound. To that end, we define a simpler Markov chain  $M'_4$ . Intuitively,  $M'_4$  is obtained by identifying states  $s_{1,k}$  and  $s_{0,k+1}$  for each  $k < N-2$ . Since, for each  $k = 1, \dots, N-2$ , the Markov chain  $M_4$  has the property that

$$\mathbb{P}[s_{1,k-1} \rightarrow s_{1,k}] = \mathbb{P}[s_{0,k} \rightarrow s_{1,k}] + \mathbb{P}[s_{0,k} \rightarrow s_{0,k+1}],$$

the transition probabilities to the left are equal for each pair of identified states. To get an upper bound for  $\tau_4$ , for the transition probabilities to the right we take the larger of  $\mathbb{P}[s_{1,k} \rightarrow s_{0,k}]$  and  $\mathbb{P}[s_{0,k} \rightarrow s_{0,k-1}]$ . Formally, we define  $M'_4$  as in Fig. A7(b). Note that we have defined  $\mathbb{P}[s_{0,0} \rightarrow s_{0,1}] = N$ , such that  $M'_4$  has only one absorbing state.

Next let us show that  $\tau_4 \leq \tau'_4$ . By construction, this is clear for traces that don't reach state  $s_{0,0}$ . For those that do, we need to show that the expected number  $s$  of steps to reach  $s_{1,N-1}$  from  $s_{0,0}$  in  $M'_4$  satisfies  $s \geq \tau_4$ . This follows from  $\mathbb{P}[s_{0,0} \rightarrow s_{0,1}]$  in  $M'_4$  is equal to  $\mathbb{P}[s_{1,N-2} \rightarrow s_{1,N-1}] + \mathbb{P}[s_{1,N-2} \rightarrow s_{0,N-2}]$  in  $M_4$ : once the trace in  $M'_4$  transitioned from  $s_{0,0}$ , the coupled trace in  $M_4$  that started at  $s_{1,N-2}$  either converged directly or transitioned to  $s_{0,N-2}$  which is to the left of  $s_{0,0}$ .

Finally, we compute  $\tau'_4$  using Eq. [23] again:

$$\begin{aligned} \tau'_4 &= N(N-1) \left( \frac{1}{N-1} + \frac{1}{2!(N-2)} + \frac{1}{3!(N-3)} + \dots + \frac{1}{(N/2)!(N-N/2)} + \underbrace{X_1 + X_2 + \dots + X_{N/2}}_{X_i \leq \frac{1}{(N/2)!}} \right) \\ &\leq \frac{N(N-1)}{N/2} \left( 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{(N/2)!} \right) + \underbrace{N(N-1) \cdot \frac{N}{2} \cdot \frac{1}{(N/2)!}}_{o(1)} \\ &\leq 2(N-1) \cdot (e-1) + o(1) = \Theta(N). \end{aligned}$$

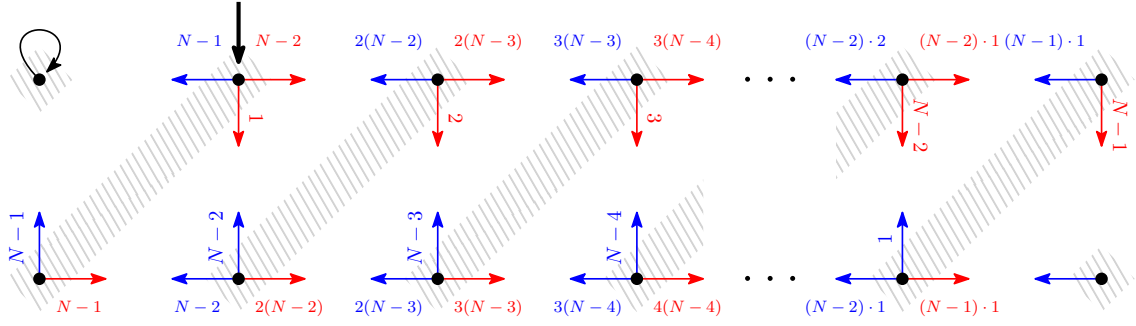
where we used that each of the last  $N/2$  terms in the first line is less than  $1/(N/2)!$  and  $\sum_{i=0}^{\infty} 1/i! = e$ .

Recovery analysis for  $L_6 = L_8$ . The Markov chain  $M_6 = M_8$  is depicted in Fig. A8.

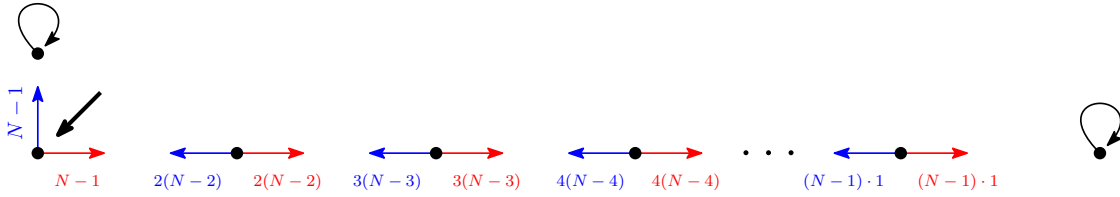
- **Recovery probability:** Identifying the states  $s_{1,k}$  and  $s_{0,k+1}$  for  $k < N-1$ , we obtain an equivalent Markov chain  $M'_6$  depicted in Fig. A9. Proposition 4 applied to  $M'_6$  immediately implies

$$\rho_6 = 1 - \frac{1}{\underbrace{1 + 1 + \dots + 1}_{N \times}} = 1 - \frac{1}{N}.$$

- **Recovery time:** For  $M'_6$  we observe that  $p_k^+ = p_k^- = \frac{k(N-k)}{N(N-1)}$ . Thus, up to rescaling, the Markov chain  $M'_6$  is equivalent to the Moran process under neutral drift (see e.g. 16, Section 4). Denoting by  $t^a$  the expected time until the process in  $M'_6$  reaches an absorbing state, and by  $t^b$  the conditional time given



**Figure A8:** Markov chain  $M_6 = M_8$ . The transition probabilities are normalized by  $N(N-1)$ . Self-loops in non-absorbing states are not shown.



**Figure A9:** Markov chain  $M_6'$  that is equivalent to  $M_6 = M_8$ . The transition probabilities are normalized by  $N(N-1)$ . We have  $\rho_6 = \rho_6'$  and  $\tau_6 = \tau_6'$ .

that it reaches the non-recovering state  $s_{0,0}$ , equations [18] and [19] from (16) rescale to

$$t^a = (N-1) \cdot H_{N-1} \quad \text{and} \quad t^b = (N-1)^2.$$

The probability of absorbing at  $s_{1,N-1}$  is  $\rho_6 = 1 - \frac{1}{N}$ . Therefore,  $t^a = (1 - \frac{1}{N}) \cdot \tau_6 + \frac{1}{N} \cdot t^b$ , or

$$\tau_6 = \frac{N \cdot t^a - t^b}{N-1} = N \cdot H_{N-1} - (N-1) = N \cdot H_N - N,$$

as desired. □

## MATLAB scripts

In the following we provide the MATLAB scripts that have been used to generate the data shown in Figs. 1–4 of the main text, and Figs. S1–S9 in this Supporting Information.

### Simulating the reputation dynamics for players with given strategies.

```
function [X,M,MEnd]=SimReputationDynamics(AssRule,ActRule,PopComp,ep,q,nIt);
% Simulates the reputation dynamics as described in Section 1.2.
```



```

% X, Matrix containing the estimated pairwise cooperation frequencies
% M, Estimated average image matrix
% MEnd, Image Matrix in the end of the simulation run.
% AssRule, Matrix that contains all assessment rules present in population
% ActRule, Matrix that contains all action rules present in the population
% PopComp, Vector that specifies how many players use each strategy
% ep, Constant error probability to commit a perception error
% q, Constant probability to observe a third-party interaction
% nIt, number of iterations of the indirect reciprocity game.

%% Setting up the objects
N=sum(PopComp); % N, Size of population
nS=length(PopComp); % nS, Number of different strategies
X=zeros(nS,nS); M=zeros(N,N); MEnd=zeros(N,N); % Initializing the output
MC=ones(N,N); % Current image matrix; initially everyone is good
CP=zeros(N,N); % Matrix that counts all individual cooperation decisions
IN=zeros(N,N); % Matrix that counts all pairwise interactions
xP=ones(1,PopComp(1)); for j=2:nS, xP=[xP, j*ones(1,PopComp(j))]; end
    % N-dim vector. The i-th entry is the strategy index of player i.
nCt=nIt/2; % initial transient time that is ignored for computing averages

%% Simulating the interactions
for t=1:nIt
    %% Choosing a donor and a recipient and letting them interact
    Do=randi(N); % Selecting a donor
    Re=Do; while Re==Do, Re=randi(N); end % Selecting a different receiver
    aD=ActRule(xP(Do),:); % Getting the donor's action rule
    stD=MC(Do,Do); stR=MC(Do,Re); % Defining the players' standings
    iA=(stD==0)*2+(stR==0)*1+1; cp=(rand(1)<aD(iA)); % cp=1 if donor cooperates

    %% Updating the donor's reputation
    for Obs=1:N % Going through all all individuals as potential observers
        if Obs==Do | Obs==Re | rand(1)<q % If individual observes interaction
            sO=AssRule(xP(Obs),:); % Getting the observer's assessment rule
            stD=MC(Obs,Do); stR=MC(Obs,Re); % Retrieving the players' standings
            if rand(1)<=1-ep % Image update without perception error
                iAs=4*(cp==0)+2*(stD==0)+1*(stR==0)+1; MC(Obs,Do)=sO(iAs);
            else % Image update with perception error
                iAs=4*(cp==1)+2*(stD==0)+1*(stR==0)+1; MC(Obs,Do)=sO(iAs);
            end
        end
    end
end
end

```

```

%% Updating the output matrices
if t>nCt% After the initial transient time has passed
    CP(Do,Re)=CP(Do,Re)+cp; IN(Do,Re)=IN(Do,Re)+1; % Updating CP and IN
    if t==nCt+1 % Updating M
        M=MC;
    else
        M=(t-nCt-1)/(t-nCt)*M+1/(t-nCt)*MC;
    end
end
end

%% Calculating output variables by averaging over all players with same strategy
MEnd=MC; % Image matrix in the end is the last current image matrix
Ms=M; M=-ones(nS,nS); % Storing M in Ms, re-initializing M
for i=1:nS % Going through all possible strategies
    for j=1:nS
        nD=find(xP==i); nR=find(xP==j); % Identifying the respective players
        if i~=j % Updating when the two strategies differ
            M(i,j)=mean(mean(Ms(nD,nR)));
            X(i,j)=sum(sum(CP(nD,nR)))/sum(sum(IN(nD,nR)));
        elseif PopComp(i)>1 % Updating within one strategy, excluding self-interactions
            M(i,j)=(sum(sum(Ms(nD,nR)))-sum(diag(Ms(nD,nR))))/(PopComp(i)*(PopComp(i)-1));
            X(i,j)=sum(sum(CP(nD,nR)))/sum(sum(IN(nD,nR)));
        end
    end
end
end
end

```

Calculating the selection-mutation equilibrium of the evolutionary dynamics for k=3 different strategies.

```

function [SMEq,Coop,AvFr,Pop]=CalcSelMutEquilibrium(b,c,mu,s,nr);
% Calculates the selection-mutation equilibrium as specified in Section 2.2
% Input parameters: b benefit of cooperation, c cost of cooperation,
% mu mutation rate, s strength of selection
% nr is a number in {1,...,8}, specifying which leading-eight strategy is considered.
% Output: SMEq vector that gives frequency of each population state in equilibrium.
% Coop Resulting average cooperation rate over course of the game
% AvFr Abundance of each of the three strategies over course of the game
% Pop List of possible population states

% Considers the competition between ALLC, ALLD and an L8-strategy Lnr
% Requires as input how often each of theses strategies would...
% cooperate against each other, as estimated by SimReputationDynamics

```

```

% This program assumes these estimates are stored in the data file 'Data L[nr].mat'
% This file contains the matrix Pop and the matrix CP.
% The matrix Pop has (N+2)(N+1)/2 rows and 3 columns; each row (n1,n2,n3)...
% gives a possible composition of the population.
% The matrix CP has (N+2)(N+1)/2 rows and 9 columns. Each row again corresponds...
% to a possible population composition. The first three columns represent how...
% often the L[nr] strategy cooperates against each other strategy; columns 4-6...
% represent the corresponding frequencies for ALLC and columns 7-9 give the...
% corresponding values for ALLD.

%% Preparations
DatName=['Data L',num2str(nr),'.mat']; load(DatName); % Loading simulation data.
nS=size(Pop,1); % Number of population states.
N=max(Pop(:,1)); % Population size.
[Pay,Coop]=CalcPay(Pop,CP,b,c,nS,N); % Calculates respective payoffs and...
    % cooperation rates in each possible population state.

%% Create the transition matrix W as specified in Section 2.2
W=zeros(nS,nS);
for iP=1:nS % Run over all possible previous states
    n=Pop(iP,:); % Previous population composition, n=[nL,nC,nD]

    % Going through all reachable population compositions
    wmp0= n(1)/N * (mu/2+(1-mu)*n(2)/(N-1)/(1+exp(-s*(Pay(iP,2)-Pay(iP,1)))));
    % Minus one L-player, plus one C-player
    wm0p= n(1)/N * (mu/2+(1-mu)*n(3)/(N-1)/(1+exp(-s*(Pay(iP,3)-Pay(iP,1)))));
    % Minus one L-player, plus one D-player
    wpm0= n(2)/N * (mu/2+(1-mu)*n(1)/(N-1)/(1+exp(-s*(Pay(iP,1)-Pay(iP,2)))));
    % Minus one C-player, plus one L-player
    w0mp= n(2)/N * (mu/2+(1-mu)*n(3)/(N-1)/(1+exp(-s*(Pay(iP,3)-Pay(iP,2)))));
    % Minus one C-player, plus one D-player
    wp0m= n(3)/N * (mu/2+(1-mu)*n(1)/(N-1)/(1+exp(-s*(Pay(iP,1)-Pay(iP,3)))));
    % Minus one D-player, plus one L-player
    w0pm= n(3)/N * (mu/2+(1-mu)*n(2)/(N-1)/(1+exp(-s*(Pay(iP,2)-Pay(iP,3)))));
    % Minus one D-player, plus one C-player

    % Inserting the values into the transition matrix
    if n(1)>0 % if the number of L[nr] individuals can further decrease
        np=[n(1)-1,n(2)+1,n(3)]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
        W(iP,iN)=wmp0;
        np=[n(1)-1,n(2),n(3)+1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
        W(iP,iN)=wm0p;
    end
    if n(2)>0 % if the number of ALLC individuals can further decrease

```

```

    np=[n(1)+1,n(2)-1,n(3)]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=wpm0;
    np=[n(1), n(2)-1,n(3)+1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=w0mp;
end
if n(3)>0 % if the number of ALLD individuals can further decrease
    np=[n(1)+1,n(2),n(3)-1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=wpm0;
    np=[n(1),n(2)+1,n(3)-1]; iN=find(Pop(:,1)==np(1)&Pop(:,2)==np(2)&Pop(:,3)==np(3));
    W(iP,iN)=w0pm;
end
W(iP,iP)=1-wmp0-wm0p-wpm0-w0mp-w0pm-w0pm;
end

v=null(W'-eye(nS,nS)); % Calculating left eigenvector of W with respect to EV 1
SMEq=v/sum(v); % Selection-mutation equilibrium is the normalized vector v
Coop=Coop*SMEq; % Average cooperation rates across all population states
AvFr=Pop'*SMEq/N; % Average abundance of each strategy across all states
end

function [Pay,Coop]=CalcPay(Pop,CP,b,c,nS,N);
% Subroutine that calculates the payoff for each of the three strategies...
% and the average cooperation rate for all possible population compositions.
% Required input: Pop, Matrix of all possible population compositions
% CP, Matrix containing the estimated pairwise cooperation frequencies
% b,c, benefit and cost of cooperation, nS, number of population states
% N, population size.

Pay=zeros(nS,3); Coop=zeros(1,nS);
for st=1:nS % going through all possible population compositions
    MC=[CP(st,1:3); CP(st,4:6); CP(st,7:9)]; % Reconstructing the image matrix
    for i=1:3, for j=1:3, % Replacing irrelevant NaN's in image matrix
        if MC(i,j)>=0; else MC(i,j)=-1; end, end
    end
    n=Pop(st,:); cp=0; % n, current population state, cp, current cooperation rate
    for iS=1:3
        if n(iS)>0
            n1=n(1)-(iS==1); n2=n(2)-(iS==2); n3=n(3)-(iS==3);
            % Remaining population from perspective of player iS
            Pay(st,iS)=(MC(1,iS)*n1+MC(2,iS)*n2+MC(3,iS)*n3)/(N-1)*b - ...
                (MC(iS,1)*n1+MC(iS,2)*n2+MC(iS,3)*n3)/(N-1)*c;
            % Estimated payoffs according to Eq.[8]
            cp=cp+n(iS)/N*(MC(iS,1)*n1+MC(iS,2)*n2+MC(iS,3)*n3)/(N-1);
            % Updating the average cooperation rate end
        end
    end
end

```

```

end
Coop(st)=cp;
end
end

```

### Calculating the selection-mutation equilibrium of the evolutionary dynamics for rare mutations.

```

function [SMEq,Coop,FXP]=CalcEqRareMutations(b,c,s,N,ep,q,nr,nIt);
% Calculates the selection-mutation equilibrium as specified in Section 2.3
% Input parameters: b benefit of cooperation, c cost of cooperation,
% s strength of selection, nr index of considered leading-eight strategy.
% N population size, ep error probability, q observation probability
% nr index of the considered leading-eight strategy % nIt number of rounds used
in SimReputationDynamics
% Output: SMEq vector for frequency of each homogeneous population in equilibrium.
% Coop Resulting average cooperation rate
% AvFr Abundance of each of the three strategies
% FXP Matrix that contains all pairwise fixation probabilities

%% Obtain the pairwise cooperation probabilities using SimReputationDynamics
AssR=[1 1 1 1 0 1 0 0;
      1 0 1 1 0 1 0 0; % Consistent Standing
      1 1 1 1 0 1 0 1; % Simple Standing
      1 1 1 0 0 1 0 1;
      1 0 1 1 0 1 0 1;
      1 0 1 0 0 1 0 1; % Stern
      1 1 1 0 0 1 0 0; % Staying
      1 0 1 0 0 1 0 0]; % Judging
% Assessment rules of leading-eight strategies
ActR=[1 0 1 1;
      1 0 1 1;
      1 0 1 0;
      1 0 1 0;
      1 0 1 0;
      1 0 1 0;
      1 0 1 0;
      1 0 1 0];
% Action rules of leading-eight strategies
AssC=[1 1 1 1 1 1 1 1]; AssD=[0 0 0 0 0 0 0 0];
ActC=[1 1 1 1]; ActD=[0 0 0 0];
% Assessment and action rules for ALLC and ALLD
[CPC,kvec]=GetPairwiseRepData([AssR(nr,:); AssC],[ActR(nr,:); ActC],N,ep,q,nIt);

```

```

    % Getting estimated cooperation probabilities when L[nr] competes with ALLC
    % for each combination [N-k k] of L[nr] and ALLC players, with k given by kvec
    [CPD,kvec]=GetPairwiseRepData([AssR(nr,:); AssD],[ActR(nr,:); ActD],N,ep,q,nIt);
    % Getting estimated cooperation probabilities when L[nr] competes with ALLD
    FXP=zeros(3,3); k=kvec';

%% Calculating pairwise fixation probabilities: L[nr] vs ALLC
PayL=(CPC(:,1).*(N-k-1)+CPC(:,3).*k)/(N-1)*b-(CPC(:,1).*(N-k-1)+CPC(:,2).*k)/(N-1)*c;
PayC=(CPC(:,2).*(N-k)+CPC(:,4).*(k-1))/(N-1)*b-(CPC(:,3).*(N-k)+CPC(:,4).*(k-1))/(N-1)*c;
    % Calculating the expected payoffs of the two strategies
RhoC=CalcFixProb(PayC,PayL,s);
RhoL=CalcFixProb(PayL(end:-1:1),PayC(end:-1:1),s);
FXP(1,2)=RhoC; FXP(2,1)=RhoL;
    % Calculating and storing fixation probabilities using a subroutine

%% Calculating pairwise fixation probabilities: L[nr] vs ALLD
PayL=(CPD(:,1).*(N-k-1)+CPD(:,3).*k)/(N-1)*b-(CPD(:,1).*(N-k-1)+CPD(:,2).*k)/(N-1)*c;
PayD=(CPD(:,2).*(N-k)+CPD(:,4).*(k-1))/(N-1)*b-(CPD(:,3).*(N-k)+CPD(:,4).*(k-1))/(N-1)*c;
    % Calculating the expected payoffs of the two strategies
RhoD=CalcFixProb(PayD,PayL,s);
RhoL=CalcFixProb(PayL(end:-1:1),PayD(end:-1:1),s);
FXP(1,3)=RhoD; FXP(3,1)=RhoL;
    % Calculating and storing fixation probabilities using a subroutine

%% Calculating pairwise fixation probabilities: ALLC vs ALLD
PayC=b*(k-1)/(N-1)-c; PayD=b*k/(N-1); % Does not require estimated payoffs
RhoC=CalcFixProb(PayC,PayD,s);
RhoD=CalcFixProb(PayD(end:-1:1),PayC(end:-1:1),s);
FXP(2,3)=RhoD; FXP(3,2)=RhoC;

%% Calculating the output
T=FXP/2; for i=1:3, T(i,i)=1-sum(T(i,:)); end % Constructing the transition matrix
v=null(T'-eye(3)); SMEq=v'/sum(v); % Calculating the selection-mutation equilibrium
CPL=SimReputationDynamics(AssR(nr,:),ActR(nr,:),N,ep,q,nIt); % Cooperation of
Leading-Eight against itself
Coop=SMEq*[CPL;1;0];
end

function pij=CalcFixProb(PayM,PayR,s);
% Calculates fixation probability according to Eq.[21]
z=exp(-s*(PayM-PayR));
pij=1/(sum(cumprod(z))+1);
end

```

```

function [CP,kvec]=GetPairwiseRepData(AssRule,ActRule,N,ep,q,nIt);
% Considers population with 2 strategies only, AssRule and ActRule have two rows
% kvec=1:N-1 counts how many players with strategy 2 are in population
% CP is an (N-1)x4 dim matrix. Each row corresponds to a population composition
% as defined through the vector k. Each entry cp_ij of a row specifies
% how often an i player would cooperate with a j player, in the following
% format: cp11, cp12, cp21, cp22.

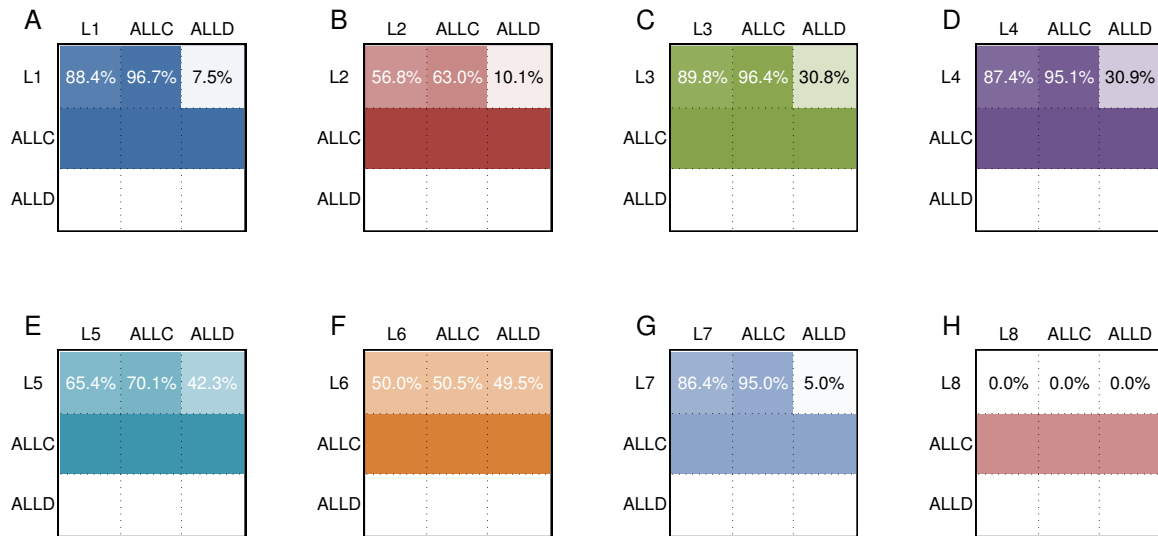
kvec=1:N-1; CP=zeros(N-1,4);
for i=1:N-1
    k=kvec(i);
    PopComp=[N-k k];
    X=SimReputationDynamics(AssRule,ActRule,PopComp,ep,q,nIt);
    CP(i,:)=[X(1,1), X(1,2), X(2,1), X(2,2)];
end
end

```

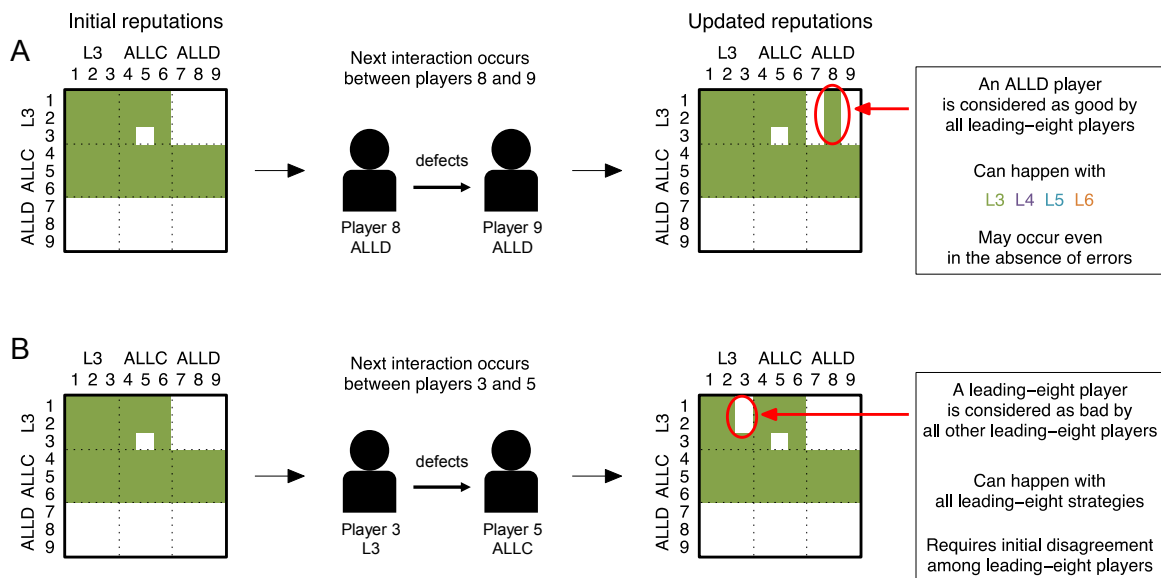
## References

- [1] Nowak, M. A. & Sigmund, K. Evolution of indirect reciprocity. *Nature* **437**, 1291–1298 (2005).
- [2] Sigmund, K. *The Calculus of Selfishness* (Princeton Univ. Press, 2010).
- [3] Sigmund, K. Moral assessment in indirect reciprocity. *Journal of Theoretical Biology* **299**, 25–30 (2012).
- [4] Ohtsuki, H. & Iwasa, Y. The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology* **239**, 435–444 (2006).
- [5] Ohtsuki, H. & Iwasa, Y. How should we define goodness? – Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology* **231**, 107–20 (2004).
- [6] Traulsen, A., Nowak, M. A. & Pacheco, J. M. Stochastic dynamics of invasion and fixation. *Physical Review E* **74**, 011909 (2006).
- [7] Szabó, G. & Tóke, C. Evolutionary Prisoner’s Dilemma game on a square lattice. *Physical Review E* **58**, 69–73 (1998).
- [8] Wu, B., Bauer, B., Galla, T. & Traulsen, A. Fitness-based models and pairwise comparison models of evolutionary games are typically different—even in unstructured populations. *New Journal of Physics* **17**, 023043 (2015).
- [9] Fudenberg, D. & Imhof, L. A. Imitation processes with small mutations. *Journal of Economic Theory* **131**, 251–262 (2006).
- [10] Imhof, L. A., Fudenberg, D. & Nowak, M. A. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences USA* **102**, 10797–10800 (2005).
- [11] Wu, B., Gokhale, C. S., Wang, L. & Traulsen, A. How small are small mutation rates? *Journal of Mathematical Biology* **64**, 803–827 (2012).
- [12] Nowak, M. A., Sasaki, A., Taylor, C. & Fudenberg, D. Emergence of cooperation and evolutionary stability in finite populations. *Nature* **428**, 646–650 (2004).
- [13] Traulsen, A. & Hauert, C. Stochastic evolutionary game dynamics. In Schuster, H. G. (ed.) *Reviews of Nonlinear Dynamics and Complexity*, 25–61 (Wiley-VCH, Weinheim, 2009).
- [14] Karlin, S. & Taylor, H. M. A. *A First Course in Stochastic Processes* (Academic, London, 1975), 2nd edition edn.
- [15] Cormen, T. H. *Introduction to algorithms* (MIT Press, Cambridge, MA, 2009).
- [16] Altrock, P. M. & Traulsen, A. Fixation times in evolutionary games under weak selection. *New Journal of Physics* **11**, 013012 (2009).

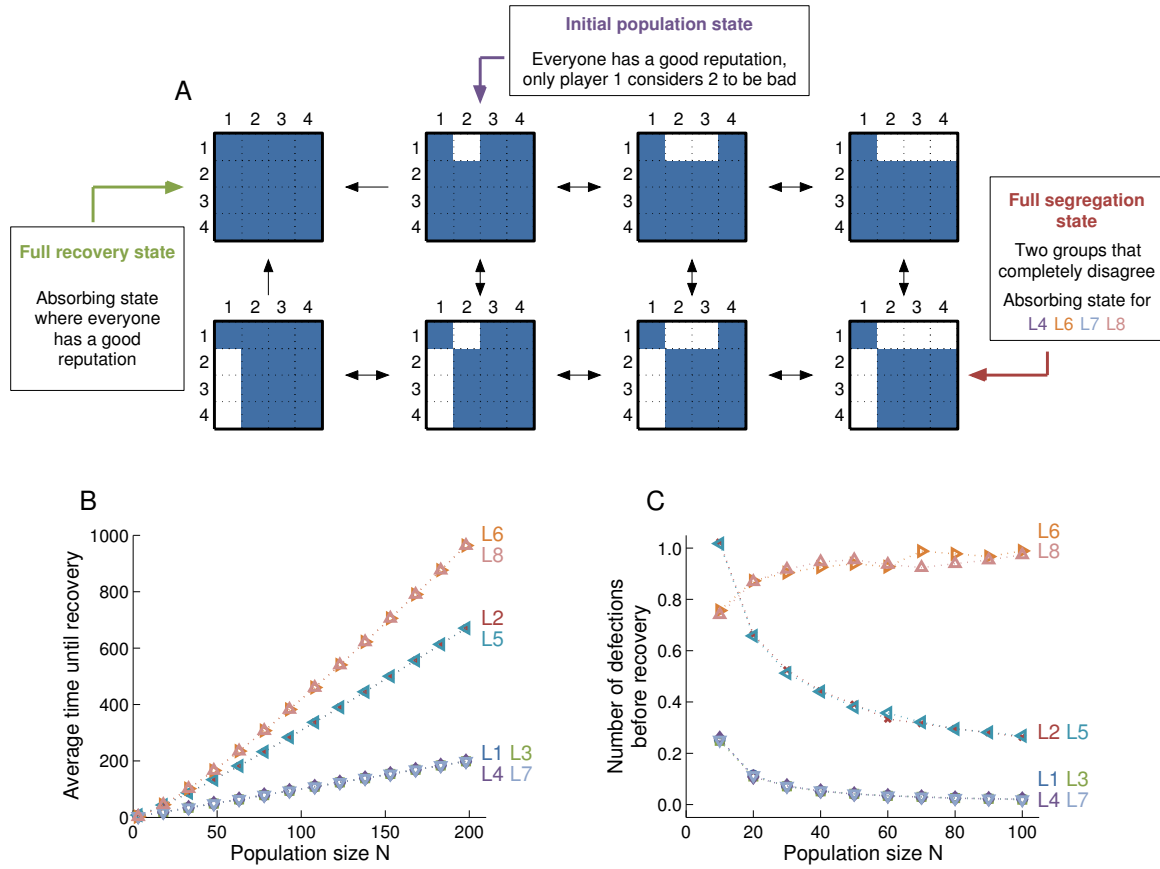




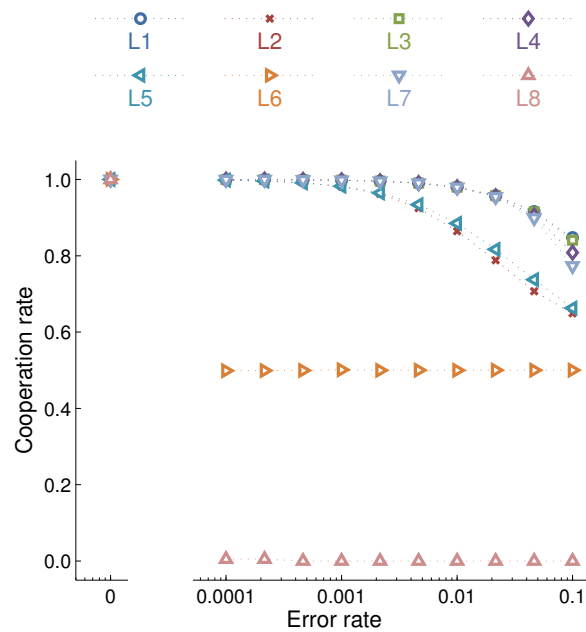
**Figure S1: The average images players have of each other converge over time.** While **Fig. 2** provides a snapshot of the reputation dynamics at a particular point in time, here we show how often players consider any other player as good on average. While ALLC players deem everyone as good and ALLD players deem everyone as bad, independent of the error rate, the leading-eight strategies differ in their assessments. Typically, a leading-eight player is most likely to consider ALLC players as good, followed by other leading-eight players and by ALLD. Whether the leading-eight strategy can resist invasion by ALLC or ALLD depends on the differences in the average image, and on the cost and benefit of cooperation. This figure uses the same parameters as in **Fig. 2**. The depicted numbers represent average values over the second half of  $2 \cdot 10^6$  rounds of the game. Results for all strategies but L7 are robust with respect to starting from a different initial image matrix.



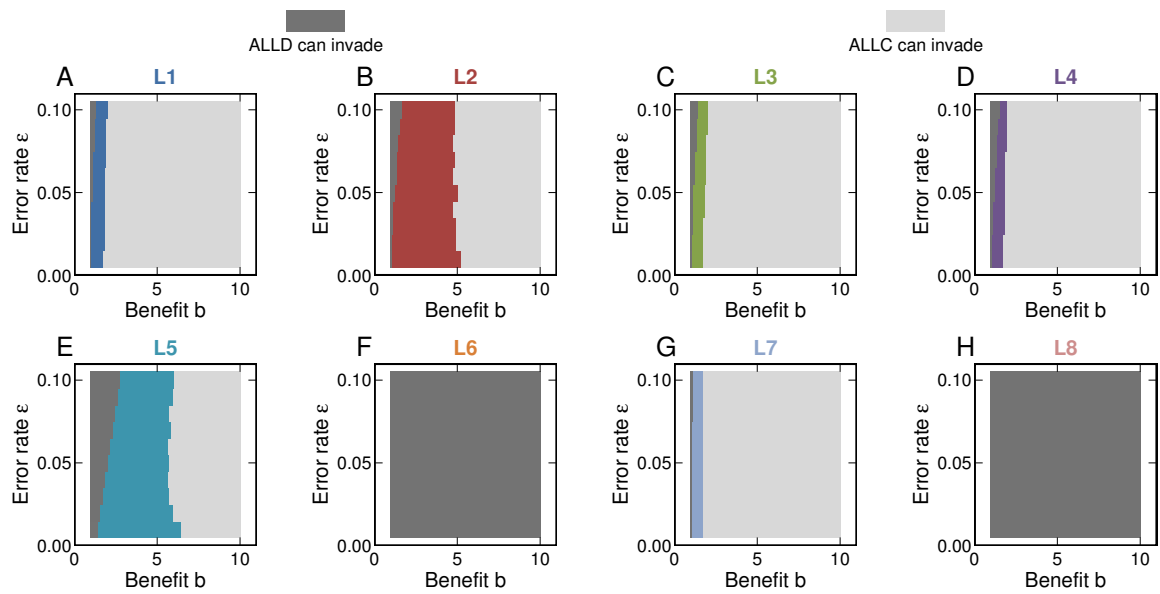
**Figure S2: The emergence of apparent misjudgments among leading-eight players.** As the game proceeds, it may occur that an ALLD player is considered as good by all leading-eight players (A), or that a leading-eight player is considered as bad by all other leading-eight players (B). Here we present stylized situations in which these cases occur. (A) For half of the leading-eight strategies (L3 – L6), an ALLD player can acquire a good reputation by defecting against another ALLD player. (B) When information is private and noisy, leading-eight strategies may disagree on the current reputation of a co-player. Here we show a case in which initially, player 3 is the only leading-eight player who deems player 5 as bad. If these two players are chosen for the next interaction, with player 3 as the donor, player 3 defects and acquires a bad reputation among her fellow leading-eight strategists.



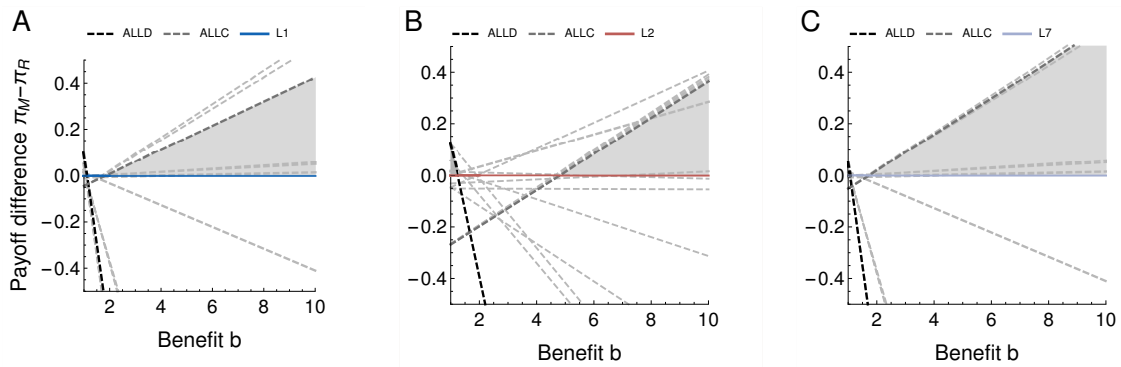
**Figure S3: The leading-eight strategies differ in their ability to recover from a single disagreements.** (A) We consider a homogeneous population in which everyone applies the same leading-eight strategy. Initially, all players are assumed to have a good reputation; only player 1 considers player 2 as bad. We derive analytical results for the reputation dynamics under the assumption that no more perception errors occur. We say the population recovers from a single disagreement, if it reaches the state where all players have a good reputation again. This is an absorbing state for all leading-eight strategies. We are interested in how likely the population recovers, and how long it takes until recovery. We find that only for four of the eight strategies, recovery is guaranteed (see SI Section 1.3). (B) With respect to the time until recovery, there are three cases. For four of the leading-eight strategies, recovery occurs quickly, and the recovery time is approximately linear. For all other strategies, including L2, we show that recovery may take substantially more time, being of order  $N \log N$ . (C) Also with respect to the expected number of defections until recovery, we observe three different qualitative behaviors. For four of the leading-eight strategies, a single perception error typically triggers no further defections, provided the population is sufficiently large. On the other extreme, for L6 and L8 we observe that any initial disagreement triggers on average one further defection.



**Figure S4: Even rare perception errors render full cooperation impossible under “Stern” (L6) or “Judging” (L8).** We have simulated how often individuals cooperate if everyone in the population applies the same leading-eight strategy. The emerging cooperation rates depend on which leading-eight strategy is chosen and how often perception errors occur. For L6 and L8 we find that even rare perception errors undermine cooperation. Among the other six strategies, L2 and L5 are more susceptible to the noise introduced by perception errors than the other four strategies. Simulations were run with a population size of  $N = 50$ , assuming complete observation,  $q = 1$ .

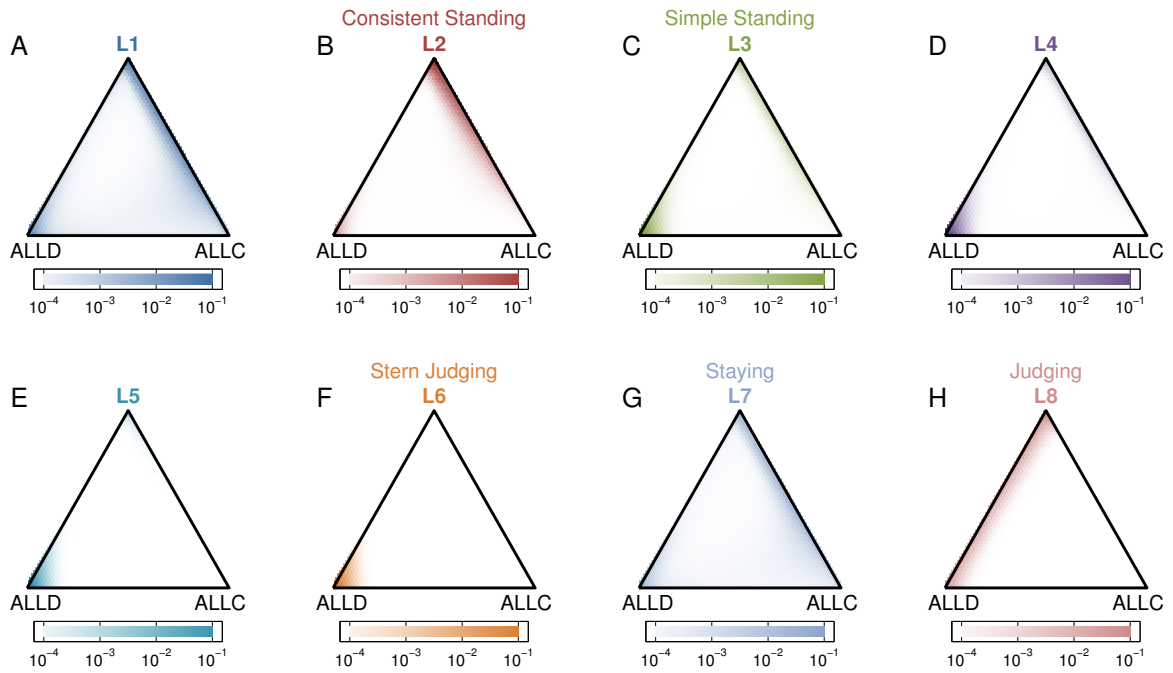


**Figure S5: Stability of each leading-eight strategy against invasion by ALLC and ALLD.** We consider a population of  $N - 1$  players adopting one of the leading-eight strategies, and we explore whether either a single ALLD mutant or an ALLC mutant gains at least the payoff of the residents. To this end, we vary the benefit of cooperation ( $x$ -axis) and the probability of perception errors ( $y$ -axis). Parameter regions in which ALLD can invade are depicted in dark grey, whereas parameter regions in which ALLC can invade are shown as light grey. Only in the colored region, the respective leading-eight strategy is stable against invasion by either ALLC or ALLD. Except for L2 and L5, we find that most leading-eight strategies are either unstable (F and H), or they only resist invasion in a small subset of the parameter space (A,C,D,G). Parameters:  $N = 50$ ,  $c = 1$ ,  $q = 0.9$ .

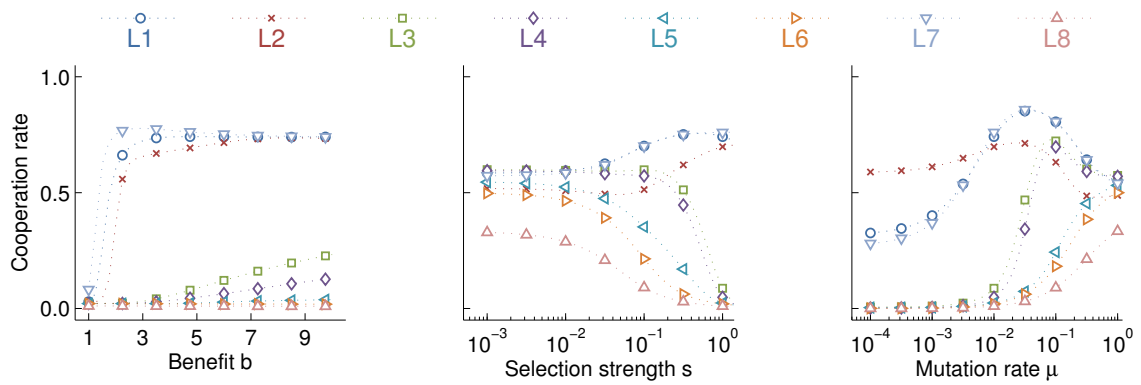


**Figure S6: Stability against ALLC and ALLD does not imply stability against all mutant invasions.**

We consider a resident population that either applies the leading-eight strategy L1 (A), L2 (B), or L7 (C). For each resident population, we consider 14 possible mutant strategies, including ALLD (black), ALLC (dark grey) and the twelve strategies that differ from the resident in only one bit (light grey). For different benefit values, we plot the payoff advantage  $\pi_M - \pi_R$  of a single mutant among  $N - 1$  residents. If for a given  $b$ -value there is a line in the upper half of the panel, the resident strategy can be invaded by the respective mutant. We find that even in the parameter region where a leading-eight strategy can resist invasion by ALLC and ALLD, other mutant strategies may be able to invade. For example, L2 can resist invasion by ALLD for  $b \gtrsim 1.5$ , and invasion by ALLC if  $b \lesssim 5$ . In between, for  $1.5 \leq b \leq 5$  there is a different mutant strategy that can invade. This mutant coincides with L2, except that it assesses a good donor who defects against a bad recipient as bad. Similar cases of successful mutants different from ALLC and ALLD also exist for L1 and L7. Parameters:  $N = 50$ ,  $c = 1$ ,  $q = 0.9$ ,  $\varepsilon = 0.05$ .

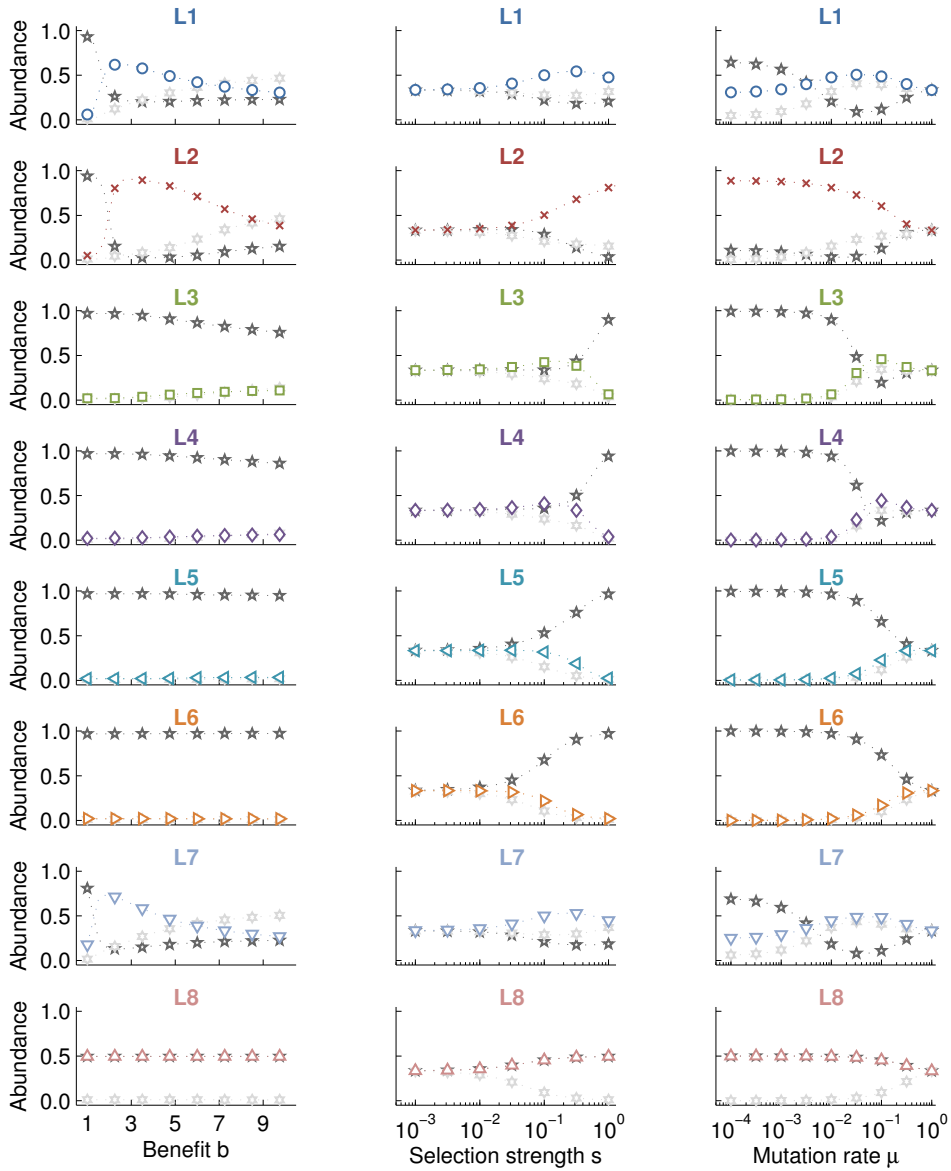


**Figure S7: When mutations are sufficiently frequent, three of the leading-eight strategies can maintain cooperation in coexistence with ALLC.** In the main text we have presented evolutionary results in the limit of rare mutations. In this limit, populations are homogeneous most of the time, rendering stable coexistences of multiple strategies impossible. Here, we present results for non-vanishing mutation rates when players can choose between a leading-eight strategy, ALLC, and ALLD. The possible population compositions are represented by a triangle. The corners of the triangle correspond to homogeneous populations, whereas interior points yield the corresponding mixed populations. The colors indicate how often the respective region of the state space is visited by the evolutionary process. (A,B,G) For the considered parameter values, we find that in three cases, a stable coexistence between a leading-eight strategy and ALLC can maintain cooperation for a considerable fraction of time. (C–F) In four cases, we find that populations typically find themselves in the vicinity of ALLD. (H) In the presence of noise, there is neutral drift between L8 and ALLD. Along that edge, no player cooperates. Parameters are the same as in **Fig. 3**, but with a strictly positive mutation rate  $\mu = 0.01$ .



**Figure S8: Cooperation is most likely to evolve for high benefit-to-cost ratios and when mutations are sufficiently frequent.** To explore the robustness of our findings, we have systematically varied three key parameters of our model, the benefit  $b$ , the selection strength  $s$ , and the mutation rate  $\mu$ . For each combination of parameter values, we have simulated the evolutionary process between ALLC, ALLD, and  $L_i$ , for each of the leading-eight strategies  $L_i$ . The figure shows the resulting cooperation rates. We recover that for usual  $b/c$ -ratios, sufficiently strong selection, and rare mutations, there are only three leading-eight strategies that can maintain some cooperation. Surprisingly, the maximum amount of cooperation is achieved for intermediate mutation rates,  $0.01 \leq \mu \leq 0.1$ . Here we observe relatively stable coexistences between ALLC and either L1 or L7. Baseline parameters:  $b=5$ ,  $s=1$  and  $\mu=0.01$ ; all other parameters are the same as in **Fig. S7**.





**Figure S9: Evolutionary abundance of the leading-eight strategies across different parameter regimes.** This figure considers the same scenario as Fig. S8, but it depicts the abundance of each strategy in the selection-mutation equilibrium. The abundance of ALLC is depicted in light grey, ALLD is shown in dark grey, and for the leading-eight strategy we use the respective color. In three cases (for L1, L2, L7) we observe that ALLC is played with positive frequency even as the selection strength  $s$  increases. In these cases, ALLC typically coexists with a majority of players who apply the leading-eight strategy.