

RESEARCH ARTICLE

Indirect punishment can outperform direct punishment in promoting cooperation in structured populations

Yujia Wen¹, Zhixue He², Chen Shen^{3*}, Jun Tanimoto^{1,3}

1 Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Fukuoka, Japan,

2 School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China,

3 Faculty of Engineering Sciences, Kyushu University, Fukuoka, Japan

* steven_shen91@hotmail.com



OPEN ACCESS

Citation: Wen Y, He Z, Shen C, Tanimoto J (2025) Indirect punishment can outperform direct punishment in promoting cooperation in structured populations. *PLoS Comput Biol* 21(6): e1013068.

<https://doi.org/10.1371/journal.pcbi.1013068>

Editor: Feng Fu, Dartmouth College, UNITED STATES OF AMERICA

Received: October 02, 2024

Accepted: April 22, 2025

Published: June 2, 2025

Copyright: © 2025 Wen et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The simulation code used to obtain the results of this work is openly available on GitHub: <https://github.com/Yujia-WEN/Indirect-reward-outperform-direct-punishment>.

Funding: This work was supported by the Japan Society for the Promotion of Science (JSPS) through the Grant-in-Aid for Scientific Research (KAKENHI), under Grant Numbers JP23H03499 (to C.S. and J.T.) and JP20H02314 (to J.T.), as well as by the

Abstract

Indirect punishment traditionally sustains cooperation in social systems through reputation or norms, often by reducing defectors' payoffs indirectly. In this study, we redefine indirect punishment for structured populations as a spatially explicit mechanism, where individuals on a square lattice target second-order defectors—those harming their neighbors—rather than their own immediate defectors, guided by the principle: “I help you by punishing those who defect against you”. Using evolutionary simulations, we compare this adapted indirect punishment to direct punishment, where individuals punish immediate defectors. Results show that within a narrow range of low punishment costs and fines, adapted indirect punishment outperforms direct punishment in promoting cooperation. However, outside this cost-fine region, outcomes vary: direct punishment may excel, both may be equally effective, or neither improves cooperation, depending on the parameter values. These findings hold even when network reciprocity alone does not support cooperation. Notably, when adapted indirect punishment outperforms direct punishment in promoting cooperation, defectors face stricter penalties without appreciably increasing punishers' costs, making it more efficient than direct punishment. Overall, our findings provide insights into the role of indirect punishment in structured populations and highlight its importance in understanding the evolution of cooperation.

Author summary

Punishment is often considered a key mechanism for maintaining cooperation in human and biological systems. Traditionally, indirect punishment relies on reputation or social norms, often by withholding cooperation, whereas we adapt this concept to penalize second-order defectors in a spatial setting. In our study, individuals on a square lattice punish those who harm their neighbors, rather than their own direct defectors. Through evolutionary simulations, we compare this adapted indirect punishment strategy with

Kakihara Science and Technology Foundation (FY2017–2018) and the Pfizer Health Research Foundation (FY2024–2025), both to J.T. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

direct punishment strategy, in which individuals punish their own immediate defective neighbors. We find that under certain conditions, adapted indirect punishment is more effective at sustaining cooperation while incurring lower costs for punishers. These results enhance our understanding of how different forms of punishment influence cooperation and offer insights into their optimal use in resolving social dilemmas.

Introduction

The advancement of society relies on cooperation among individuals to maximize collective benefits [1–4]. However, cooperation imposes individual costs while benefiting others, making it vulnerable to exploitation by selfish free-riders [5]. Understanding how cooperation emerges among strangers, particularly in one-shot and anonymous encounters [6], is crucial because these settings artificially remove key mechanisms that typically sustain cooperation. Such mechanisms include repeated interactions [7–10], reputation effects [11,12], and other reciprocity mechanisms [13], creating the most challenging conditions for cooperation to emerge. Yet, despite these obstacles, economic game experiments consistently show that people do cooperate, with cooperation levels stabilizing at lower but persistent rates [14,15]. Evidence suggests that cooperation in such scenarios may result from prosocial preferences or confusion about the game structure [16–19]. Nevertheless, researchers continue to explore mechanisms to sustain cooperation, particularly through costly punishment [6,20]. Beyond human interactions, punishment is widely observed in nature [21,22]. In animal populations, individuals often use punishment to constrain non-cooperative or deceptive behaviors, thereby promoting population stability [23]. These findings suggest that punishment is a fundamental enforcement mechanism shaping cooperative behavior across biological and social systems.

Costly punishment generally takes two forms: direct punishment and indirect punishment. Direct punishment occurs when an individual penalizes a defector who has personally harmed them [20,24–26]. While this strategy can deter defection, it imposes a cost on the punishers and risks retaliation, making it evolutionarily unstable unless supported by additional mechanisms such as reputation effects [27], voluntary participation [28], or network reciprocity [29–35]. By contrast, indirect punishment occurs when an individual punishes a defector who has harmed someone else, even though the punisher was not personally affected. This enforcement mechanism helps maintain cooperation by discouraging selfish behavior on a broader social scale. A well-documented example is indirect reciprocity, where individuals refuse to cooperate with those who have previously defected against someone else [12,36–38]. Experimental studies suggest that even when indirect punishment is rarely used, its mere availability can sustain cooperative behavior by reinforcing prosocial norms [37]. Unlike direct punishment, which requires personal retaliation, indirect punishment typically relies on social mechanisms such as reputation systems to track defectors' past actions and coordinate sanctions [11,12,38]. While it may lower retaliation risks, it still incurs costs, such as monitoring effort and the risk of misjudging defectors.

Owing to these differences, the evolutionary dynamics of direct punishment have been extensively studied in structured populations, where network structures stabilize it by balancing second-order free-riding (cooperators who don't punish) and antisocial punishment (targeting cooperators), thus ensuring their survival [29,30,32]. By contrast, indirect punishment has received less attention in spatial networks, as it traditionally relies on reputation or norms to sanction defectors across broader social scales, often among strangers [37,38]. In structured populations where interactions are localized, direct punishment can naturally be applied to

nearby defectors. However, monitoring reputations and enforcing norms can be challenging, potentially weakening the effectiveness of aforementioned indirect punishment within such populations. Despite these constraints, the fundamental principle of indirect punishment—sanctioning defectors for others' sake—suggests that alternative spatially adapted mechanisms, independent of global reputation, may still facilitate cooperation. This contrast raises a pivotal question: in structured populations, how does spatially adapted indirect punishment compare to direct punishment in sustaining cooperation, and under what conditions might it prove more effective?

To address this gap, we examined the effectiveness of direct and indirect punishment in structured populations using an extended evolutionary prisoner's dilemma game on a square lattice. The game consists of two stages: in the first stage, players decide whether to cooperate with their direct neighbors, and in the second stage, they choose whether to impose fines on defectors at a personal cost. Direct punishment is implemented by allowing players to punish their immediate neighbors (depicted as the purple-shaded area in Fig 1). Meanwhile, indirect punishment is adapted to the square lattice, requiring focal players to punish their indirect neighbors, specifically the neighbors of their neighbors (illustrated as the cyan-shaded area in Fig 1). Although this adaptation differs from traditional reputation-based indirect punishment models, it maintains the core idea of indirect punishment, where defectors are sanctioned even if they have not directly harmed the punisher.

Through evolutionary simulations, we show that compared to scenarios without punishment, indirect punishment promotes cooperation more effectively than direct punishment under conditions of low punishment costs and fines. Outside this region, when costs are low but fines are high, both forms of punishment are equally effective. When both the cost and fine are relatively high, direct punishment becomes more favorable. However, in other cases, neither punishment strategy significantly enhances cooperation. These results are robust in terms of when network reciprocity alone can and cannot sustain cooperation. In scenarios where indirect punishment outperforms direct punishment, defectors encounter stricter

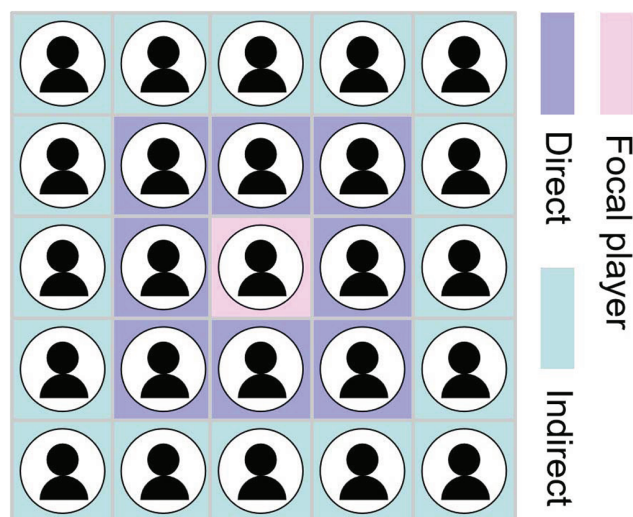


Fig 1. Schematic diagram of direct and indirect punishment in the model. In the first stage of the prisoner's dilemma game, a focal player interacts with their eight direct neighbors (purple-shaded area). In the second stage, punishment targets differ: direct punishment imposes penalties on defective direct neighbors, while indirect punishment targets defective second-order neighbors (up to sixteen, cyan-shaded area), excluding direct neighbors.

<https://doi.org/10.1371/journal.pcbi.1013068.g001>

penalties without substantially increasing the costs borne by punishers, making it a more efficient enforcement mechanism. These findings suggest that under specific low-cost and low-fine conditions, indirect punishment serves as a more effective method for promoting cooperation than direct punishment, offering valuable insights into optimizing punishment mechanisms within structured populations.

Model and method

Considering a population of size N placed on a square lattice with moore neighborhood and periodic boundaries, in which players play two-stage prisoner’s dilemma game (PDG) with their neighbors [39]. In the first stage of the PDG, paired players simultaneously choose to either cooperate or defect, with the outcomes determined by the actions of both parties. Specifically, players can obtain payoff R from mutual cooperation and a P from mutual defection. When unilateral defection occurs, cooperator receives S while defector receives T . The social dilemma arises when the payoff elements satisfy $T > R > P > S$ and $2R > T + S$, indicating that while mutual cooperation maximizes collective benefits, defection provides the highest individual benefit. To simplify the model without losing generality, we set $R = 1$ and $P = 0$, and follow ref.[40] to introduce the dilemma strength parameter $r := \frac{T-R}{R-P} = \frac{P-S}{R-P}$, which characterizes the relative benefit of defection over cooperation. Thus, $T = 1 + r$ and $S = -r$.

In the second stage, players decide whether to punish defectors based on first-stage outcomes. Punishment involves punishers incurring a cost (γ) to impose a fine (β) to others. Based on first-stage actions, we define four strategies: cooperating without punishing (CN), defecting without punishing (DN), cooperating and punishing defectors (CP), and defecting while punishing defectors (DP). The payoffs for these four strategies are summarized in Table 1.

In direct punishment, a focal player interacts with their direct neighbors in both stages (highlighted in purple in Fig 1). In contrast, indirect punishment differs in that second-stage interactions target second-order neighbors (i.e., neighbors of first-order neighbors, highlighted in cyan in Fig 1). Notably, punishing second-order defectors ultimately benefits first-order neighbors, embodying the principle: “I help you by punishing those who defect against you”. From this perspective, this form of punishment can be viewed as an indirect reward to direct neighbors.

The total payoff of a player is the sum of the outcomes from the two-stage interactions. Specifically, in the case of direct punishment, the focal player’s payoffs in the first and second stages are calculated according to Table 1(a) and 1(b), respectively, based on interactions with eight direct neighbors. In the case of indirect punishment, the focal player’s payoff

Table 1. The payoff matrix for the two-stage Prisoner’s Dilemma game. (a) The first-stage interaction involves players deciding whether to cooperate; (b) The second stage involves players deciding whether to punish, with the interaction outcomes under different strategies representing the row player’s payoff in that stage.

(a) The first-stage interaction					(b) The second-stage interaction				
	CN	DN	CP	DP		CN	DN	CP	DP
CN	1	-r	1	-r	CN	0	0	0	0
DN	1+r	0	1+r	0	DN	0	0	-β	-β
CP	1	-r	1	-r	CP	0	-γ	0	-γ
DP	1+r	0	1+r	0	DP	0	-γ	-β	-β -γ

<https://doi.org/10.1371/journal.pcbi.1013068.t001>

in the first stage is calculated through interactions with eight immediate neighbors according to Table 1(a), while the second-stage payoff is obtained through interactions with sixteen second-order neighbors according to Table 1(b).

We use Monte Carlo simulations to obtain the results for both the direct punishment and indirect reward models. Initially, players randomly adopt one of the four strategies. In each time step of the Monte Carlo simulation, N updates occur. During each update, a focal player is randomly selected, and their payoff is calculated according to the above rules, the player then updates strategy through pairwise imitation. Thus, on average, each player updates strategy once per time step. The focal player i mimics the strategy of a randomly selected direct neighbor j with a Fermi probability [41]:

$$w_{i \leftarrow j} = \frac{1}{1 + \exp\{(\Pi_i - \Pi_j)\kappa\}}, \quad (1)$$

where Π_x represents the player x 's total payoffs. The parameter κ denotes the imitation strength; as $\kappa \rightarrow 0$, imitation approaches randomness, while as $\kappa \rightarrow +\infty$, imitation is driven by payoff differences. In this study, we set $\kappa = 10$ to investigate a strong imitation scenario. The network sizes N range from 100^2 to 400^2 in simulation, and the final results are averaged from over 50 independent simulations, with each simulation producing average values from the last 2000 time steps out of a total exceeding 1×10^5 time steps.

Results

In structured populations, network reciprocity can enable cooperative clusters to resist defectors [3,42], thereby sustaining cooperation. However, under high dilemma strength, network reciprocity fails to prevent defectors from infiltrating cooperative clusters, making cooperation difficult to establish. To investigate the role of punishment in promoting cooperation, we examine two representative scenarios: low dilemma strength ($r = 0.05$), where cooperation coexists with defection at 56 % in the absence of punishment, and high dilemma strength ($r = 0.2$), where cooperation vanishes without punishment. These two scenarios serve as benchmarks for comparing the effectiveness of punishment, and our model reverts to these baselines when punishment costs substantially exceed fines.

Interestingly, indirect punishment can promote cooperation more effectively than direct punishment within a narrow region of punishment parameters. The results in Fig 2 indicate four distinct regions in the parameter space of punishment costs and fines, revealing the differences between indirect and direct punishment in promoting cooperation, regardless of dilemma strength. In region I, characterized by low-costs and low-fines relative to payoff of mutual cooperation, with the fine slightly higher than the punishment cost, indirect punishment can outperform direct punishment, as shown in the red circle of Fig 2. In region II, where both costs and fines are relatively high than that of region I, direct punishment is more effective than indirect punishment. In region III, characterized by low-costs and high-fines, both indirect and direct punishment perform similarly, enabling full cooperation, as seen in Figs A and B in S1 Appendix. In region IV, where costs typically exceed fines, neither punishment method improves cooperation significantly, and outcomes revert to those observed without punishment.

Indirect punishment are more sensitive to changes in punishment costs than direct punishment, which affects the overall efficiency of the punishment mechanism. At a certain fine level (e.g., $\beta = 0.2$), as the cost of punishment increases, it typically reduces the effectiveness of direct punishment. This is evident in the upper left of Fig 3, where the proportion of CP

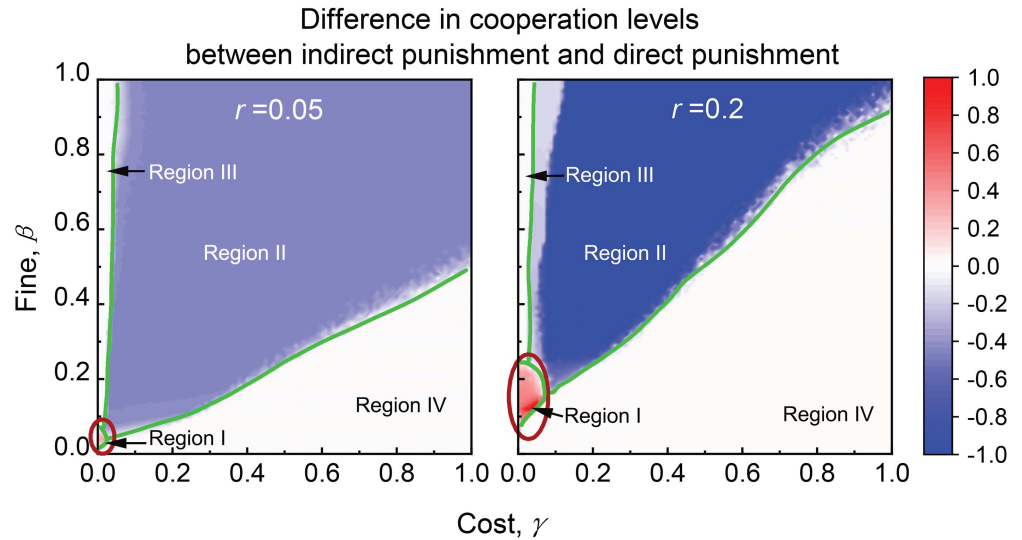


Fig 2. Indirect punishment can promote cooperation more effectively than direct punishment under conditions of low-costs and low-fines (marked with dark red circles). Color codes illustrate the difference in cooperation levels (i.e., the fraction of $CN + CP$ within the population) between direct punishment and indirect punishment as functions of punishment cost γ and fine β for low dilemma strength (left panel) and high dilemma strength (right panel). Based on the comparison of indirect punishment to direct punishment in promoting cooperation, the parameter space can be divided into the following regions: Region I, where indirect punishment is superior to direct punishment; Region II, where direct punishment is superior to indirect punishment; Region III and region IV, where the difference between direct and indirect punishment is negligible.

<https://doi.org/10.1371/journal.pcbi.1013068.g002>

gradually declines as γ increases, ultimately resulting in the collapse of cooperation. For $\gamma < 0.04$, the proportion of cooperators (i.e., $CN + CP$) under indirect punishment can exceed that under direct punishment, as shown in the lower left of Fig 3. Specifically, indirect punishment induces pure cooperation within the population for $\gamma < 0.02$. However, as γ continues to rise, the efficiency of indirect punishment declines sharply, leading to the disappearance of cooperation when $\gamma > 0.1$. In contrast, cooperation in the direct punishment scenario only disappears when $\gamma > 0.15$.

At a given punishment cost level (e.g., $\gamma = 0.03$), as the fine increases, direct punishment induces the spontaneous emergence of CP and gradually dominates the population when $\beta > 0.12$, as shown in the upper-right panel of Fig 3. In contrast, under indirect punishment, cooperation can be established when $\beta = 0.1$. However, as β increases, second-order free riders (i.e., CN , who benefit from others' punishments without bearing the cost) gradually replace CP . Cooperation reemerges only when $\beta > 0.58$. Additionally, in indirect punishment, the fine required to eliminate defection is $\beta = 0.66$, significantly higher than the $\beta = 0.26$ required for direct punishment. This suggests that indirect punishment requires a larger fine to achieve the same effect as direct punishment.

We then investigate why indirect punishment can be more or less effective than direct punishment in promoting cooperation by analyzing the characteristics of population evolution. To do so, we fix $\beta = 0.2$ and employ a block-like initial distribution to better understand the competitive dynamics between strategies, as shown in Fig 4. Under low-cost conditions (e.g., $\gamma = 0.01$), whether through direct punishment or indirect punishment, DN can easily invade CN and DP in the direct punishment scenario, as DN individuals free-ride without bearing the costs of punishment. It is evident that the proportion of CP steadily increases

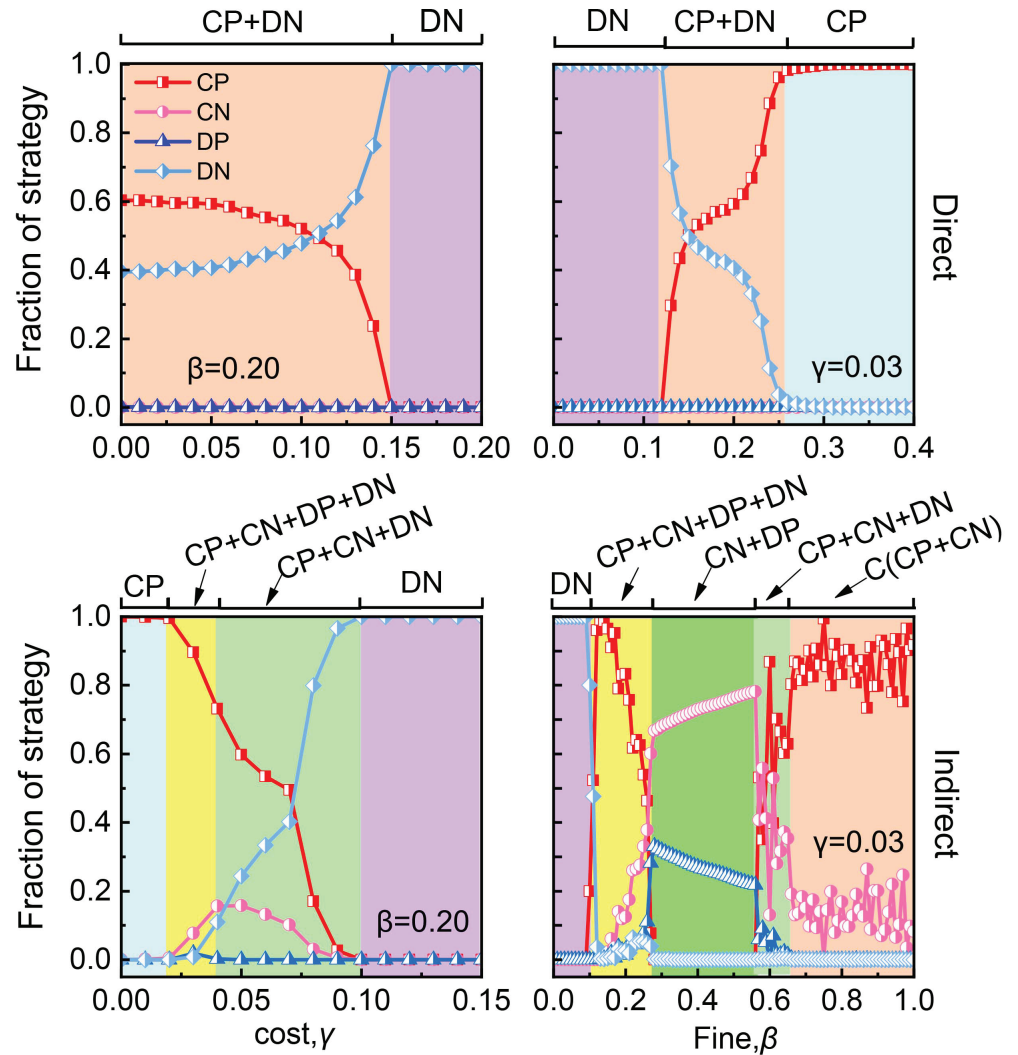


Fig 3. The effectiveness of indirect punishment in promoting cooperation is sensitive to the costs of punishment. Depicted are the fraction of *CN*, *DN*, *CP* and *DP* as functions of punishment cost γ (left panels) and fine β (right panels) for direct punishment and indirect punishment. Results are obtained under high dilemma strength. Other parameters are set to $\beta = 0.2$ for left panels, and γ for right panels.

<https://doi.org/10.1371/journal.pcbi.1013068.g003>

from the early stages of evolution, indicating that the implementation of punitive measures by *CP* allows them to resist invasion by *DN* and contributes to their dominance. Eventually, a coexistence of *CP* and *DN* emerges under direct punishment, as shown in the first row of Fig 4. In contrast, indirect punishment can support the further expansion of *CP* clusters, see the second row of Fig 4. When *DN* clusters are surrounded by *CP*, *DN* on the edge of these clusters, encounters more second-order punishment neighbors, resulting in they suffer more punishment, see the upper panels of Fig 5. Although *CP* impose more punishment under the indirect punishment, the total punishment cost for *CP* does not increase significantly due to the lower cost associated with punishment action. By increasing the frequency of punishment on defectors without substantially raising punishers' costs at low γ , indirect punishment allow *CP* to outperform *DN*. As a result, indirect punishment can be more efficient than direct punishment in promoting cooperation.

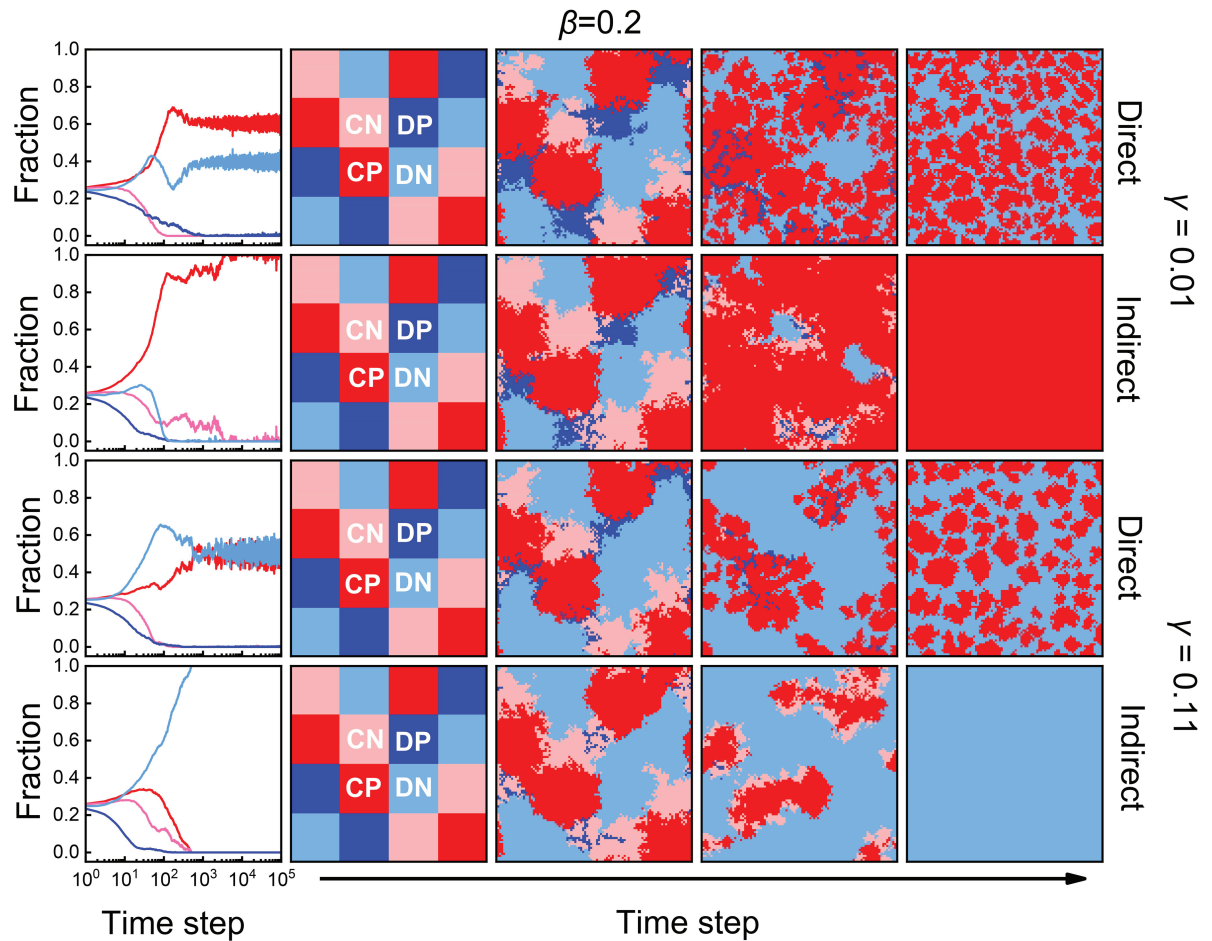


Fig 4. Indirect punishment promote the expansion of CP clusters when the punishment cost is low, but lead to their collapse when the cost becomes high. The leftmost panels illustrate the proportion of the four strategies as a function of time steps, while the right panels present typical evolutionary snapshots starting from a block-like initial distribution for direct punishment and indirect punishment. Light red, dark red, light blue, and dark blue represent CN, CP, DN, and DP, respectively. The results are obtained under fixed $\beta = 0.2$ and $r = 0.2$.

<https://doi.org/10.1371/journal.pcbi.1013068.g004>

For a high cost (e.g., $\gamma = 0.11$), the size of CP clusters shrinks under direct punishment, but CP can still resist invasion of DN, as shown in the third row of Fig 4. In contrast, indirect punishment fail to prevent the invasion of DN, as shown in the fourth row of Fig 4. When CP clusters are surrounded by DN, CP individuals have more second-order defecting neighbors, leading to an increase in punitive actions. However, the rise in the total punishment cost limits the expansion of CP. Furthermore, despite the increase in punishment imposed by CP, these punishments are not effectively concentrated on defectors at the cluster boundaries. As shown in the lower panel of Fig 5, the total fines of DN imposed do not increase significantly. Consequently, indirect punishment fail to sustain cooperation.

Conclusion and discussion

To conclude, in this study, we introduced a novel form of indirect punishment on a square lattice, where players impose penalties solely on defective second-order neighbors (the neighbors of their direct neighbors). This spatially adapted strategy indirectly benefits direct neighbors through a spillover effect, enhancing cooperation within the network. Our results

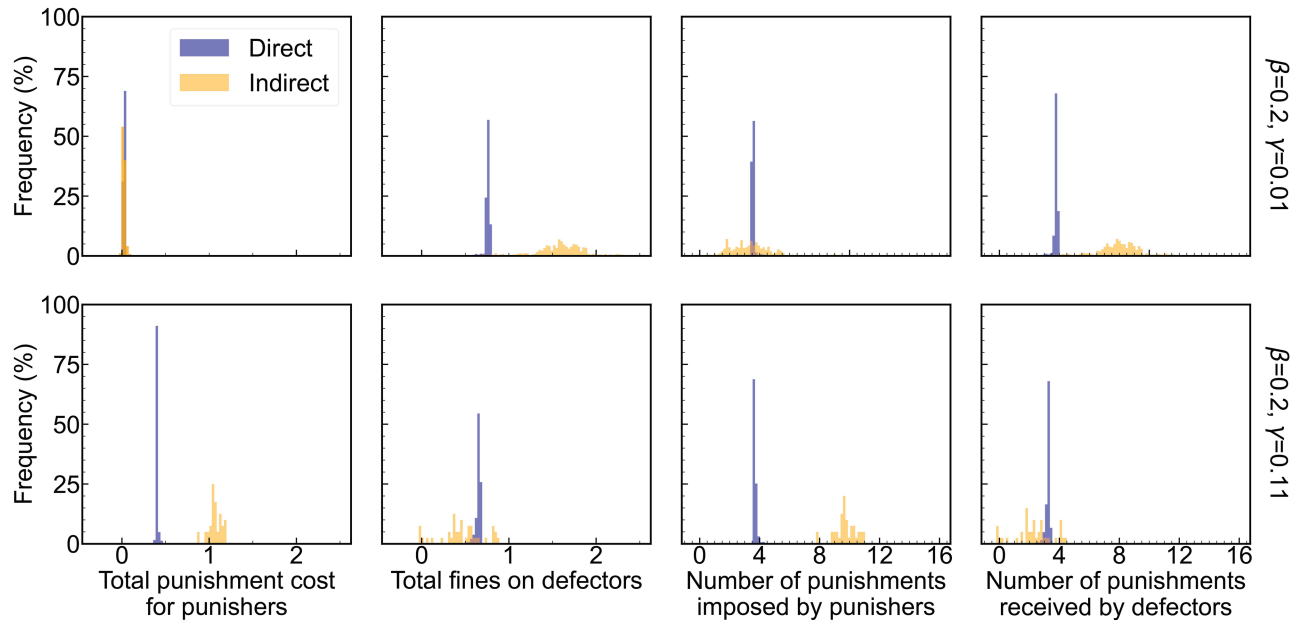


Fig 5. Indirect punishment behavior can lead to defectors being punished more without significantly increasing the total cost for punishers when the punishment cost is low, making them more effective than direct punishment in promoting cooperation. However, high punishment costs result in a greater total punishment cost in the case of indirect punishment, which diminishes the payoff advantage for punishers, making them less effective than direct punishment. Depicted are the frequency distributions of the total punishment cost for punishers (first column panels from the left), the total fines imposed on the defectors (second column panels), the number of punishments imposed by punishers (third column panels), and the number of punishments received by defectors who are punished (fourth column panels) for direct punishment and indirect punishment. To analyze the evolutionary characteristics of punishers, the distribution data include all evolutionary time steps prior to the stabilization of strategy proportions within the population, and the evolutionary processes correspond to the results presented in Fig 4.

<https://doi.org/10.1371/journal.pcbi.1013068.g005>

demonstrate that indirect punishment outperforms direct punishment—which targets defective direct neighbors—in fostering cooperation under low-cost and low-fine parameter conditions (see region I in Fig 2). In low-cost, high-fine conditions (see region III in Fig 2), both strategies are equally effective, while direct punishment proves superior under high-cost, high-fine conditions (see region II in Fig 2). These findings remain robust across different dilemma strengths, providing valuable insights into the comparative effectiveness of punishment strategies and highlighting the critical role of the cost-to-fine ratio in determining their success within structured populations.

By exploring the comparative effectiveness of direct punishment and indirect punishment, we offer new insights into how different punishment forms promote cooperation in structured populations. Most existing evolutionary models focus on direct punishment because it is unlikely to be evolutionarily stable and requires explanation. Behavioral experiments further support this, showing that individuals who achieve the highest total payoff tend not to use costly (direct) punishment, suggesting this behavior may have evolved for other reasons [20]. In contrast, a form of indirect punishment—characterized by withholding cooperation—does not face the same issues as direct punishment, such as second-order free riding. Consequently, evolutionary game literature often questions why people continue to use direct punishment when they could opt to withhold cooperation [43,44].

Moving beyond these evolutionary considerations, we identified a narrow parameter region in structured populations where indirect punishment is more effective than direct punishment. This finding is somewhat surprising because direct punishment directly undermines

defectors' payoffs at the boundaries of cooperative clusters—a critical mechanism for network reciprocity. In contrast, indirect punishment targets defective second-order neighbors, encouraging them to cooperate for the benefit of their direct neighbors, yet it does not directly alter the competitive dynamics at the cluster boundaries. These results highlight the nuanced roles that direct and indirect punishment play in fostering cooperation and suggest that their relative effectiveness is closely linked to the underlying network structure.

This unexpected effectiveness of indirect punishment behaviors, particularly under conditions of low costs and fines, prompts the need for a deep explanation of the mechanisms driving this outcome. We propose a novel explanation: under these conditions, indirect punishment leads to significant accumulated fines on defectors without significantly increasing the punishers' total costs. This is due to the broader punishment scope on indirect neighbors, allowing defectors to be penalized from multiple sources, which amplifies the overall punitive effect, as illustrated in Figs 4 and 5. In contrast, direct punishment targets defectors within a limited range, reducing its impact. This explanation deepens our understanding of how network reciprocity supports altruistic behavior, traditionally focused on the formation and maintenance of cooperative clusters [42,45,46].

In evolutionary game theory, direct punishment refers to a mechanism where an individual actively and intentionally penalizes another for their behavior, typically at a personal cost, such as reducing an immediate neighbor's payoff in a spatial lattice. Conversely, indirect punishment involves enforcing cooperation through social norms, reputation, or third-party systems, extending influence beyond direct individual action. In our adaptation for a square lattice, we redefine indirect punishment to target second-degree neighbors, capturing its broader spatial scope. While some might note that this retains elements of direct punishment—like individual costs or retaliation risks—these features stem from the lattice's localized structure and do not undermine the distinction. The critical difference lies in the punishment range: immediate for direct, extended for indirect. This approach preserves the spirit of indirect punishment while enabling a precise analysis of how punishment range influences cooperation in structured populations, enhancing the study of spatial evolutionary dynamics.

Although our adapted form of indirect punishment does not fully align with previous definition—which emphasizes withholding cooperation rather than imposing sanctions[20, 36]—our model remains valid as it captures the core concept of sanctioning defectors who harm others. In spatial structures, interactions are limited to local neighbors, prompting us to adapt a novel form of indirect punishment to assess its influence on cooperation in networked populations. Implementing the traditional definition would require higher-order networks to capture group interactions beyond the capabilities of a square lattice [47,48], as indirect punishment derived from indirect reciprocity involves ternary or higher-order relationships. Future work could explore higher-order networks to more accurately reflect the dynamics of withholding cooperation.

Additionally, extending the current model by considering broader punishment scopes beyond indirect neighbors could test the robustness of indirect punishment's effectiveness compared to direct punishment. Such extensions are meaningful because they can quantify the potential future returns for costly punishers on the effectiveness of cooperation and may reveal additional evolutionary mechanisms behind the evolution of costly punishment. Incorporating reinforcement learning algorithms [49] or large language models [50] could also provide a framework for agents to adapt their punishment strategies based on learned experiences, offering insights into how these strategies evolve over time. Finally, considering the competition between different forms of punishment is also meaningful, as the competitive outcomes may help understand various forms of punishment can evolve and mechanisms that drive preferences for punishment forms, warrant further investigation.

Supporting information

S1 Appendix. This file contains two figures, each showing the fractions of the *CN*, *DN*, *CP*, and *DP* strategies as functions of the punishment cost γ and fine β , under both direct punishment (top row) and indirect punishment (bottom row) scenarios. Figs A and B present results for low dilemma strength ($r = 0.05$) and high dilemma strength ($r = 0.2$), respectively. (PDF)

Author contributions

Conceptualization: Yujia Wen, Chen Shen.

Formal analysis: Yujia Wen, Zhixue He, Chen Shen.

Funding acquisition: Chen Shen, Jun Tanimoto.

Investigation: Yujia Wen, Chen Shen.

Methodology: Yujia Wen, Zhixue He, Chen Shen.

Project administration: Chen Shen, Jun Tanimoto.

Software: Yujia Wen, Zhixue He.

Supervision: Jun Tanimoto.

Validation: Zhixue He.

Writing – original draft: Yujia Wen, Zhixue He, Chen Shen, Jun Tanimoto.

Writing – review & editing: Chen Shen, Jun Tanimoto.

References

1. Axelrod R, Hamilton WD. The evolution of cooperation. *Science*. 1981;211(4489):1390–6. <https://doi.org/10.1126/science.7466396> PMID: 7466396
2. Ohtsuki H, Hauert C, Lieberman E, Nowak MA. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*. 2006;441(7092):502–5. <https://doi.org/10.1038/nature04605> PMID: 16724065
3. Nowak MA. Five rules for the evolution of cooperation. *Science*. 2006;314(5805):1560–3. <https://doi.org/10.1126/science.1133755> PMID: 17158317
4. Boyd R, Richerson PJ. Culture and the evolution of human cooperation. *Philos Trans R Soc Lond B Biol Sci*. 2009;364(1533):3281–8. <https://doi.org/10.1098/rstb.2009.0134> PMID: 19805434
5. Hardin G. The tragedy of the commons. *Classic papers in natural resource economics revisited*. New York: Routledge. 2018. p. 145–56.
6. Fehr E, Gächter S. Altruistic punishment in humans. *Nature*. 2002;415(6868):137–40. <https://doi.org/10.1038/415137a> PMID: 11805825
7. Camerer CF. *Behavioral game theory: experiments in strategic interaction*. Princeton University Press; 2011.
8. Hilbe C, Chatterjee K, Nowak MA. Partners and rivals in direct reciprocity. *Nat Hum Behav*. 2018;2(7):469–77. <https://doi.org/10.1038/s41562-018-0320-9> PMID: 31097794
9. Hu S, Leung C, Leung H. Modelling the dynamics of multiagent q-learning in repeated symmetric games: a mean field theoretic approach. *Adv Neural Inf Process Syst*. 2019.
10. Zhu Y, Xia C, Chen Z. Nash equilibrium in iterated multiplayer games under asynchronous best-response dynamics. *IEEE Trans Automat Contr*. 2023;68(9):5798–805. <https://doi.org/10.1109/tac.2022.3230006>
11. Xia C, Wang J, Perc M, Wang Z. Reputation and reciprocity. *Phys Life Rev*. 2023;46:8–45. <https://doi.org/10.1016/j.pprev.2023.05.002> PMID: 37244154
12. Nowak MA, Sigmund K. Evolution of indirect reciprocity by image scoring. *Nature*. 1998;393(6685):573–7. <https://doi.org/10.1038/31225> PMID: 9634232

13. Rand DG, Nowak MA. Human cooperation. *Trends Cogn Sci*. 2013;17(8):413–25. <https://doi.org/10.1016/j.tics.2013.06.003> PMID: 23856025
14. Chaudhuri A. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Exp Econ*. 2011;14:47–83.
15. Andreoni J, Varian H. Preplay contracting in the prisoners' dilemma. *Proc Natl Acad Sci U S A*. 1999;96(19):10933–8.
16. Fehr E, Fischbacher U, Gächter S. Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum Nat*. 2002;13(1):1–25. <https://doi.org/10.1007/s12110-002-1012-7> PMID: 26192593
17. Wang G, Li J, Wang W, Niu X, Wang Y. Confusion cannot explain cooperative behavior in public goods games. *Proc Natl Acad Sci U S A*. 2024;121(10):e2310109121. <https://doi.org/10.1073/pnas.2310109121> PMID: 38412126
18. Burton-Chellew MN, West SA. Payoff-based learning best explains the rate of decline in cooperation across 237 public-goods games. *Nat Hum Behav*. 2021;5(10):1330–8. <https://doi.org/10.1038/s41562-021-01107-7> PMID: 33941909
19. Shen C, He Z, Guo H, Hu S, Tanimoto J, Shi L, et al. Beyond a binary theorizing of prosociality. *Proc Natl Acad Sci U S A*. 2024;121(49):e2412195121. <https://doi.org/10.1073/pnas.2412195121> PMID: 39602256
20. Dreber A, Rand DG, Fudenberg D, Nowak MA. Winners don't punish. *Nature*. 2008;452(7185):348–51. <https://doi.org/10.1038/nature06723> PMID: 18354481
21. Clutton-Brock TH, Parker GA. Punishment in animal societies. *Nature*. 1995;373(6511):209–16. <https://doi.org/10.1038/373209a0> PMID: 7816134
22. Raihani NJ, Thornton A, Bshary R. Punishment and cooperation in nature. *Trends Ecol Evol*. 2012;27(5):288–95. <https://doi.org/10.1016/j.tree.2011.12.004> PMID: 22284810
23. Riehl C, Frederickson ME. Cheating and punishment in cooperative animal societies. *Philos Trans R Soc Lond B Biol Sci*. 2016;371(1687):20150090. <https://doi.org/10.1098/rstb.2015.0090> PMID: 26729930
24. Wu J-J, Zhang B-Y, Zhou Z-X, He Q-Q, Zheng X-D, Cressman R, et al. Costly punishment does not always increase cooperation. *Proc Natl Acad Sci U S A*. 2009;106(41):17448–51. <https://doi.org/10.1073/pnas.0905918106> PMID: 19805085
25. Li X, Jusup M, Wang Z, Li H, Shi L, Podobnik B, et al. Punishment diminishes the benefits of network reciprocity in social dilemma experiments. *Proc Natl Acad Sci U S A*. 2018;115(1):30–5. <https://doi.org/10.1073/pnas.1707505115> PMID: 29259113
26. Zhu Y, Zhang Z, Xia C, Chen Z. Equilibrium analysis and incentive-based control of the anticompeting networked game dynamics. *Automatica*. 2023;147:110707. <https://doi.org/10.1016/j.automatica.2022.110707>
27. dos Santos M, Rankin DJ, Wedekind C. The evolution of punishment through reputation. *Proc Biol Sci*. 2011;278(1704):371–7. <https://doi.org/10.1098/rspb.2010.1275> PMID: 20719773
28. Fowler JH. Altruistic punishment and the origin of cooperation. *Proc Natl Acad Sci U S A*. 2005;102(19):7047–9. <https://doi.org/10.1073/pnas.0500938102> PMID: 15857950
29. Perc M, Szolnoki A. Self-organization of punishment in structured populations. *New J Phys*. 2012;14(4):043013.
30. Szolnoki A, Perc M. Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment. *Phys Rev X*. 2017;7(4):041027. <https://doi.org/10.1103/physrevx.7.041027>
31. Guo T, He Z, Shi L. Self-organization in mobile populations promotes the evolution of altruistic punishment. *Phys A: Statist Mech Appl*. 2023;630:129282. <https://doi.org/10.1016/j.physa.2023.129282>
32. Helbing D, Szolnoki A, Perc M, Szabó G. Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput Biol*. 2010;6(4):e1000758. <https://doi.org/10.1371/journal.pcbi.1000758> PMID: 20454464
33. Fu F, Chen X. Leveraging statistical physics to improve understanding of cooperation in multiplex networks. *New J Phys*. 2017;19(7):071002. <https://doi.org/10.1088/1367-2630/aa78c1> PMID: 29606900
34. Wang X, Fu F, Wang L. Deterministic theory of evolutionary games on temporal networks. *J R Soc Interface*. 2024;21(214):20240055. <https://doi.org/10.1098/rsif.2024.0055> PMID: 38807526
35. Chen Z, Geng Y, Chen X, Fu F. Unbending strategies shepherd cooperation and suppress extortion in spatial populations. *New J Phys*. 2024;26(7):073047.
36. Rockenbach B, Milinski M. The efficient interaction of indirect reciprocity and costly punishment. *Nature*. 2006;444(7120):718–23. <https://doi.org/10.1038/nature05229> PMID: 17151660

37. Ule A, Schram A, Riedl A, Cason TN. Indirect punishment and generosity toward strangers. *Science*. 2009;326(5960):1701–4. <https://doi.org/10.1126/science.1178883> PMID: 20019287
38. Balafoutas L, Nikiforakis N, Rockenbach B. Direct and indirect punishment among strangers in the field. *Proc Natl Acad Sci U S A*. 2014;111(45):15924–7. <https://doi.org/10.1073/pnas.1413170111> PMID: 25349390
39. Szabó G, Fath G. Evolutionary games on graphs. *Phys Rep*. 2007;446(4–6):97–216.
40. Wang Z, Kokubo S, Jusup M, Tanimoto J. Universal scaling for the dilemma strength in evolutionary games. *Phys Life Rev*. 2015;14:1–30. <https://doi.org/10.1016/j.plrev.2015.04.033> PMID: 25979121
41. Sigmund K, De Silva H, Traulsen A, Hauert C. Social learning promotes institutions for governing the commons. *Nature*. 2010;466(7308):861–3. <https://doi.org/10.1038/nature09203> PMID: 20631710
42. Wang Z, Kokubo S, Tanimoto J, Fukuda E, Shigaki K. Insight into the so-called spatial reciprocity. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2013;88(4):042145. <https://doi.org/10.1103/PhysRevE.88.042145> PMID: 24229153
43. Ohtsuki H, Iwasa Y, Nowak MA. Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature*. 2009;457(7225):79–82. <https://doi.org/10.1038/nature07601> PMID: 19122640
44. Panchanathan K, Boyd R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*. 2004;432(7016):499–502. <https://doi.org/10.1038/nature02978> PMID: 15565153
45. Nowak MA, May RM. The spatial dilemmas of evolution. *Int J Bifurcation Chaos*. 1993;03(01):35–78. <https://doi.org/10.1142/s0218127493000040>
46. Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. Statistical physics of human cooperation. *Phys. Rep*. 2017;687:1–51. <https://doi.org/10.1016/j.physrep.2017.05.004>
47. Guo H, Jia D, Sendiña-Nadal I, Zhang M, Wang Z, Li X, et al. Evolutionary games on simplicial complexes. *Chaos Solitons Fract*. 2021;150:111103. <https://doi.org/10.1016/j.chaos.2021.111103>
48. Civilini A, Sadekar O, Battiston F, Gómez-Gardeñes J, Latora V. Explosive cooperation in social dilemmas on higher-order networks. *Phys Rev Lett*. 2024;132(16):167401. <https://doi.org/10.1103/PhysRevLett.132.167401> PMID: 38701463
49. Hu S, Soh H, Piliouras G. The best of both worlds in network population games: reaching consensus and convergence to equilibrium. *Adv Neural Inf Process Syst*. 2024.
50. Ren S, Cui Z, Song R, Wang Z, Hu S. Emergence of social norms in generative agent societies: principles and architecture. In: *Proceedings of the 33rd International Joint Conference on Artificial Intelligence*. 2024.