SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

January 2025, Vol. 68, Iss. 1, 112204:1–112204:14
https://doi.org/10.1007/s11432-023-4170-3

# Exit options sustain altruistic punishment and decrease the second-order free-riders, but it is not a panacea

Chen SHEN[1,2], Zhao SONG[3], Lei SHI[2*], Jun TANIMOTO[1] & Zhen WANG[4*]

[1]*Faculty of Engineering Sciences, Kyushu University, Fukuoka 816-8580, Japan;*
[2]*School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming 650221, China;*
[3]*School of Computing, Engineering and Digital Technologies, Teesside University, Middlesbrough TS1, UK;*
[4]*School of Artificial Intelligence, OPtics and ElectroNics (iOPEN), Northwestern Polytechnical University,*
*Xi'an 710072, China*

Received 14 August 2023/Revised 20 June 2024/Accepted 4 September 2024/Published online 26 December 2024

**Abstract** Why do humans punish free riders at their own expense? This behavior represents an evolutionary puzzle in human societies. This study explores the role of exit strategies in fostering altruistic punishment within evolutionary game theory. We extend the traditional prisoner's dilemma model by incorporating exiters, players who opt out for a small payoff while nullifying their opponent's payoff, and altruistic punishers who cooperate and punish non-cooperators. Our findings indicate that in well-mixed populations, exiters destabilize defection but do not promote altruistic punishment. In social networks, however, exiters enable altruistic punishment via cyclic dominance among altruistic punishers, defectors, and exiters. Notably, this cyclic dominance is sensitive to exit payoffs; adjustments can lead to cyclic dominance of non-punishing cooperators, defectors, and exiters, or to a bi-stable state between these two types of cyclic dominance. These results highlight the nuanced impact of exiters on altruistic punishment, emphasizing the need for careful incentivization of exit behavior. While exiters can support altruistic punishment in networked populations, their effectiveness is not a panacea and is highly sensitive to exit payoffs, indicating limits to the voluntary participation mechanism.

**Keywords** evolutionary game theory, altruistic punishment, coexistence, cyclic dominance, bi-stable

**Citation** Shen C, Song Z, Shi L, et al. Exit options sustain altruistic punishment and decrease the second-order free-riders, but it is not a panacea. Sci China Inf Sci, 2025, 68(1): 112204, https://doi.org/10.1007/s11432-023-4170-3

## 1 Introduction

In recent years, artificial intelligence (AI) has opened new frontiers in human-AI cooperation, promising transformative impacts across various domains. Successful human-AI interactions rely on principles of human-human cooperation, such as trust, empathy, and effective communication, which inform the design of AI systems. Understanding these dynamics is essential for developing AI that integrates seamlessly into human social structures [1–3]. However, cooperation presents a puzzle: it often incurs individual costs, while the temptation to free-ride-benefit without contributing-threatens to undermine it. According to evolutionary principles, notably the 'survival of the fittest', such free-riding behavior could have evolutionary advantages, potentially leading to the breakdown of cooperation [4].

To build AI systems that fit smoothly into human social contexts, understanding the complexities of cooperation is crucial. Evolutionary game theory identifies several mechanisms [5, 6] to explain and foster cooperation, including direct, indirect, and network reciprocity, as well as kin and group selection. However, in one-shot and anonymous games where players cannot repeatedly play with the same opponent or lack information about their opponents, these mechanisms are absent. In such scenarios, cooperation relies on social mechanisms like altruistic punishment [7,8], rewards [9–11], voluntary participation [12,13] and among others [14–16].

While social mechanisms like altruistic punishment are crucial in these scenarios, they come with several limitations. Altruistic punishment, where cooperators punish defectors, can promote cooperation

---

[*] Corresponding author (email: shi_lei65@hotmail.com, w-zhen@nwpu.edu.cn)

but relies on identifying and tracking non-cooperators, which is challenging in reality [13]. Additionally, it can lead to retaliation, where defectors punish cooperators, deterring individuals from punishing non-cooperators [17]. Moreover, punishment is costly, and non-punishing cooperators who avoid these costs fare better than altruistic punishers. If non-punishing cooperators dominate, defectors can eventually overtake the population [17, 18]. These limitations pose significant challenges to the effectiveness of altruistic punishment, making it crucial to understand its emergence alongside cooperation. For a comprehensive review, interested readers are recommended to refer to existing literature on this topic [19, 20].

Contrary to altruistic punishment, voluntary participation, traditionally represented by loners, offers a different solution to the cooperation conundrum. Loners abstain from public goods, opting out and receiving a fixed payoff without influencing public goods. This mechanism can establish cooperation in one-shot and anonymous scenarios through cyclic dominance: cooperators give way to defectors, who give way to loners, who in turn give way to cooperators. Unlike altruistic punishment, it avoids the complexities of tracking defectors and associated evolutionary challenges. Studies indicate that it can even address the second-order free rider problem (those who cooperate but do not punish) and elucidate the emergence of altruistic punishment [21–23]. In an infinite population, evolutionary dynamics may lead to a Nash equilibrium involving punishing and non-punishing cooperators or an oscillating state among defectors, non-punishing cooperators, and loners [21, 22]. Moreover, the efficacy of loners extends to finite populations, showcasing their versatility in promoting cooperative outcomes [23].

One intriguing variant of the loner strategy is represented by exiters [24]. Unlike loners, exiters not only exit the game for a small positive payoff but also nullify their opponent's payoff to zero. This can be exemplified in collaborative scientific projects: if one researcher free-rides or abandons the project, the remaining researcher's chances of success diminish. Although this strategy harms both cooperators and defectors in one-shot games, it can support cooperation through repeated interactions, reputation, or social networks.

Building on studies that employ the voluntary participation mechanism to explain altruistic punishment, it is natural to extend this concept to include exiters. By examining the impact of exiters on altruistic punishment, we aim to determine the extent to which exiters can explain this behavior. This exploration contributes to the robustness of the voluntary participation mechanism in explaining altruistic punishment, as exiters, though slightly different from loners, still fall within the voluntary participation concept.

To achieve this, we extend the traditional prisoner's dilemma model by incorporating exiters and altruistic punishers alongside traditional cooperators and defectors. Traditional cooperators are non-punishing cooperators who benefit others at a personal cost without punishing defectors. Traditional defectors do nothing, while altruistic punishers benefit others at a personal cost and punish defectors. Exiters simply exit the game, nullifying their opponent's payoff. We consider two population structures: a well-mixed population, where players have an equal probability of interacting with others, and a regular lattice, where players only interact with their nearest neighbors. Our findings reveal that in well-mixed populations (finite or infinite), exiters destabilize defection but do not promote the emergence of altruistic punishment. However, in social networks, exiters promote altruistic punishment through cyclic dominance among altruistic punishers, defectors, and exiters. This cyclic dominance is not always stable; adjustments to the exit payoff can replace it with cyclic dominance of non-punishing cooperators, defectors, and exiters or a bi-stable state between these two types of cyclic dominance. These results indicate that exiters have a nuanced effect on altruistic punishment, highlighting the sensitivity of outcomes to exit payoffs and population structures and the need to carefully incentivize exit behavior to encourage altruistic punishment.

## 2 Methods

We studied the evolution of altruistic punishment in a two-stage prisoner's dilemma game by introducing two other action types, altruistic punishment and exit. In the first stage, each individual must make a choice simultaneously between cooperation (C), defection (D), and exit (E). In the second stage, cooperators decide whether to punish the defectors at a personal cost to themselves $\gamma$. To the defectors, this means an imposed fine $\beta$. This process results in four possible actions.

- AP, cooperate and punish defectors. Those who cooperate and punish are altruistic punishers because they punish free riders even at the expense of their interests.

**Table 1** Payoff matrix for the weak prisoner's dilemma game with altruistic punishment and an exit option.

| | AP | NC | D | E |
|---|---|---|---|---|
| AP | 1 | 1 | $-\gamma$ | 0 |
| NC | 1 | 1 | 0 | 0 |
| D | $b-\beta$ | $b$ | 0 | 0 |
| E | $\epsilon$ | $\epsilon$ | $\epsilon$ | $\epsilon$ |

- NC, cooperate but do not punish defectors. These non-punishing cooperators are also known as second-order free riders because free-riding on punishment saves the cost of punishing the defectors.
- D, defect but do not punish. These are also known as first-order free riders.
- E, exit the game in favor of a small but positive payoff $\epsilon$ irrespective of whom they encounter. They do not participate in these two stages.

In a typical prisoner's dilemma game, mutual cooperation (defection) generates the reward (punishment) $R$ ($P$). If one player cooperates and the other defects, the cooperative player gets the sucker's payoff $S$, and the defected player obtains the temptation to defect $T$. For simplicity, we chose the weak prisoner's dilemma game as our base model by setting $R = 1$, $P = S = 0$, and $T = b$. To ensure that the weak prisoner's dilemma game satisfies the payoff ranking of the strict prisoner's dilemma game ($T > R > P > S$), we set $1 \leqslant b < 2$. Additionally, we set $\epsilon < 1$ to ensure that cooperating yields a higher payoff than exiting, reflecting real-world situations where cooperation is more beneficial than opting out.

The parameters are summarized in Table 1. Furthermore, we held the cost of punishment $\gamma$ and the fine for defectors constant at 0.1 and 0.3, respectively. These fixed settings do not limit our conclusion to specific cases. By varying the temptation value $b$, we can explore the effects of the exit option on altruistic punishment in different scenarios. The difference $b-\beta$ relative to 1 is crucial: if $b-\beta > 1$, defectors gain a payoff advantage, potentially increasing defection. Conversely, if $b-\beta \leqslant 1$, altruistic punishers and cooperators gain a competitive edge, enhancing cooperation.

### 2.1 Finite population

We first considered a finite and well-mixed population of $N$ individuals. Each individual adopted the Moran process [25], also known as a frequently dependent process, to select their action. At each time step, a randomly selected player $i$ with fitness $f_i = \mathrm{e}^{s\Pi_i}$ ($\Pi_i$ is the actual payoff of the individual $i$ obtained through their interaction) updates its action by imitating the action of player $j$ with fitness $f_j = \mathrm{e}^{s\Pi_j}$ who is selected with a probability proportional to its fitness. Here, $s$ is the selection strength, the condition of $s \to 0$ corresponds to the weak selection and evolution proceeds as a neutral drift. We further assumed that with a small probability $\mu$, players randomly select their action from the rest of the possible actions, otherwise, the imitation process is governed by the Moran process. When the mutation rate is small ($\mu \to 0$), the population will either eliminate the mutant or allow it to fully invade the population, resulting in homogeneity with one dominant actor most of the time.

Suppose that there are only two actions in the population, i.e., actions A and B, and these actions can be one of the four actions among the full action set $\{a, b, c, d\}$. Here, the symbols $a, b, c, d$ represent AP, NC, D, and E, respectively. In a finite population of size $N$ with $x$ A and $y = N - x$ B actions, the average payoff of $\Pi_{xy}$ and $\Pi_{yx}$ to players with A and B actions is the following:

$$\begin{aligned}\Pi_{\mathrm{AB}} &= \frac{(x-1)P_{\mathrm{AA}}+(N-x)P_{\mathrm{AB}}}{N-1},\\ \Pi_{\mathrm{BA}} &= \frac{xP_{\mathrm{BA}}+(N-x-1)P_{\mathrm{BB}}}{N-1},\end{aligned} \tag{1}$$

where $P_{\mathrm{AB}}$ is the payoff obtained from the single encounter of actors A and B, and so are payoffs $P_{\mathrm{AA}}$, $P_{\mathrm{BA}}$, and $P_{\mathrm{BB}}$. This allows us to describe the evolutionary dynamics of the population in terms of a Markov Chain of size 4 [26–28]. Given the above assumptions, the probability of changing the number of $x$ individuals with action A in a population of $y = N - x$ individuals with action B by $\pm 1$, $T^{\pm}_{\mathrm{AB}}$, is

$$\begin{aligned}T^{+}_{\mathrm{AB}} &= \frac{xf_i}{xf_i+yf_j}\frac{y}{N},\\ T^{-}_{\mathrm{AB}} &= \frac{yf_j}{xf_i+yf_j}\frac{x}{N},\end{aligned} \tag{2}$$

and hence the fixation probability $\rho_{\mathrm{AB}}$ of a single mutant actor A within a population of $N-1$ B actors can be derived as [29, 30]

$$\rho_{\mathrm{AB}}=\frac{1}{\sum_{k=0}^{N-1}\prod_{x=1}^{k}\frac{T_{\mathrm{AB}}^{-}}{T_{\mathrm{AB}}^{+}}}=\frac{1}{\sum_{k=0}^{N-1}\prod_{x=1}^{k}\frac{\mathrm{e}^{s\Pi_{\mathrm{BA}}}}{\mathrm{e}^{s\Pi_{\mathrm{AB}}}}}. \tag{3}$$

The fixation probabilities $\rho_{\mathrm{AB}}$ define the transition probabilities of the a Markov process between four different homogeneous states of the population, with the following associated transition matrix:

$$\begin{array}{c} \\ \mathrm{AP} \\ \mathrm{NC} \\ \mathrm{D} \\ \mathrm{E} \end{array}\begin{array}{c} \begin{array}{cccc} \mathrm{AP} & \mathrm{NC} & \mathrm{D} & \mathrm{E} \end{array} \\ \begin{pmatrix} \rho_{aa} & \frac{\mu\rho_{ab}}{3} & \frac{\mu\rho_{ac}}{3} & \frac{\mu\rho_{ad}}{3} \\ \frac{\mu\rho_{ba}}{3} & \rho_{bb} & \frac{\mu\rho_{bc}}{3} & \frac{\mu\rho_{bd}}{3} \\ \frac{\mu\rho_{ca}}{3} & \frac{\mu\rho_{cb}}{3} & \rho_{cc} & \frac{\mu\rho_{cd}}{3} \\ \frac{\mu\rho_{da}}{3} & \frac{\mu\rho_{db}}{3} & \frac{\mu\rho_{dc}}{3} & \rho_{dd} \end{pmatrix} \end{array}, \tag{4}$$

where $\rho_{AA}=1-\sum_{A\neq B}\frac{\mu\rho_{AB}}{3}, A, B\in\{a,b,c,d\}$. The stationary distribution of each strategy, which is determined by the normalized right eigenvector to the largest eigenvalue, provides both the relative evolutionary advantage of each strategy and the stationary fraction of actors, as discussed in sources such as [23, 31]. For any pair of strategies A and B in the finite population, natural selection favors B replacing A only if $\rho_{AB}>\frac{1}{N}$ [29].

### 2.2 Infinite population

We then employed replicator dynamics to analyze the evolutionary outcomes in an infinite and well-mixed population. Let $x, y, z, w$ denote the fractions of altruistic punishers (AP), non-punishing cooperators (NC), defectors (D), and exiters (E) in the population, where $0\leqslant x, y, z, w\leqslant 1$ and $x+y+z+w=1$. The replicator equations are

$$\begin{aligned} \dot{x}&=x\left(\Pi_{\mathrm{AP}}-\overline{\Pi}\right), \\ \dot{y}&=y\left(\Pi_{\mathrm{NC}}-\overline{\Pi}\right), \\ \dot{z}&=z\left(\Pi_{\mathrm{D}}-\overline{\Pi}\right), \\ \dot{w}&=w\left(\Pi_{\mathrm{E}}-\overline{\Pi}\right). \end{aligned} \tag{5}$$

The symbols $\Pi_{\mathrm{AP}}$, $\Pi_{\mathrm{NC}}$, $\Pi_{\mathrm{D}}$, and $\Pi_{\mathrm{E}}$ denote the average payoff of altruistic punishers, non-punishing cooperators, defectors, and exiters. Whereas $\overline{\Pi}=x\Pi_{\mathrm{AP}}+y\Pi_{\mathrm{NC}}+z\Pi_{\mathrm{D}}+w\Pi_{\mathrm{E}}$ is the average payoff of the whole population. According to the defined payoffs in Table 1, we obtain the following equation:

$$\begin{aligned} \Pi_{\mathrm{AP}}&=x+y-z\gamma, \\ \Pi_{\mathrm{NC}}&=x+y, \\ \Pi_{\mathrm{D}}&=x(b-\beta)+yb, \\ \Pi_{\mathrm{E}}&=\epsilon. \end{aligned} \tag{6}$$

Using the constraint $w=1-x-y-z$, we obtain

$$\begin{cases} \dot{x}=f\left(x,y,z\right)=x\left[\left(1-x\right)\left(\Pi_{\mathrm{AP}}-\Pi_{\mathrm{E}}\right)-y\left(\Pi_{\mathrm{NC}}-\Pi_{\mathrm{E}}\right)-z\left(\Pi_{\mathrm{D}}-\Pi_{\mathrm{E}}\right)\right], \\ \dot{y}=g\left(x,y,z\right)=y\left[\left(1-y\right)\left(\Pi_{\mathrm{NC}}-\Pi_{\mathrm{E}}\right)-x\left(\Pi_{\mathrm{AP}}-\Pi_{\mathrm{E}}\right)-z\left(\Pi_{\mathrm{D}}-\Pi_{\mathrm{E}}\right)\right], \\ \dot{z}=h\left(x,y,z\right)=z\left[\left(1-z\right)\left(\Pi_{\mathrm{D}}-\Pi_{\mathrm{E}}\right)-y\left(\Pi_{\mathrm{NC}}-\Pi_{\mathrm{E}}\right)-x\left(\Pi_{\mathrm{AP}}-\Pi_{\mathrm{E}}\right)\right]. \end{cases} \tag{7}$$

For the detailed stability analysis of each equilibrium, please refer to Appendix A in the supplementary information.

### 2.3 Networked population

Different from well-mixed populations, global interactions in which an individual can interact with any other individual are no longer possible in the networked population. Instead, networks only allow local interactions, which means that individuals can only interact with their direct neighbors. Our basic network structure is a two dimensional regular lattice with periodic boundary conditions. Each node was occupied by one individual, and each individual could only interact with its neighbors along its links. Our

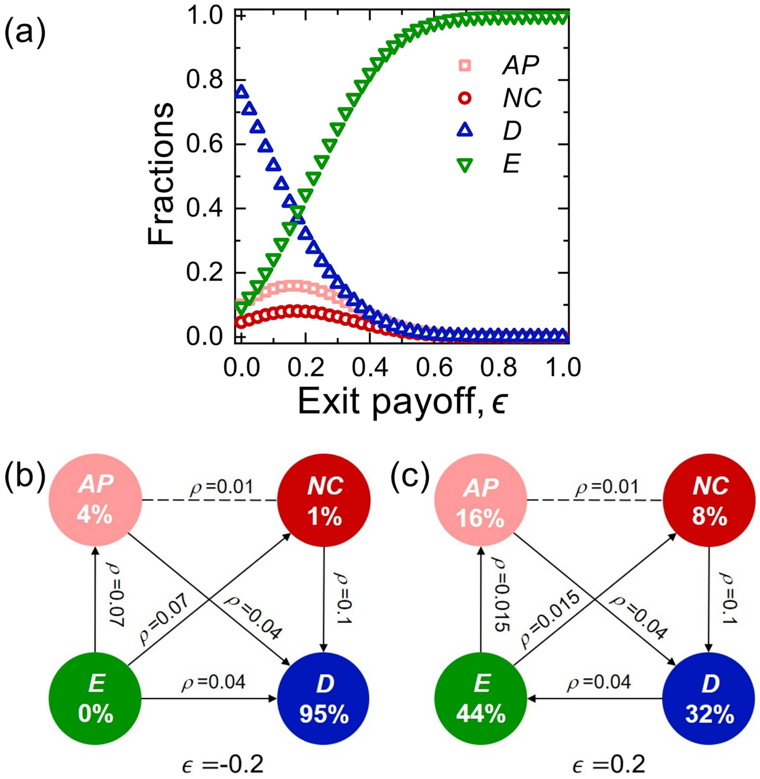

**Figure 1** (Color online) Exiters establish altruistic punishment in a finite population, but altruistic punishers struggle to dominate the population. (a) Stationary probability distributions of each actors independence on the exit payoff $\epsilon$; (b) transition probabilities for each pair of actors when the exit payoff is negative; (c) transition probabilities for each pair of actors when the exit payoff is positive. The parameter values are $b = 1.5$, $\beta = 0.3$, $\gamma = 0.1$, $s = 0.2$, $N = 100$.

simulation contained the following steps. Initially, each individual was designed as either an altruistic punisher (AP), a non-punishing cooperator (NC), a defector (D), or an exiter (E) with equal probability. Each player acquires their total payoff by playing with all their direct neighbors according to the payoff matrix defined in Table 1. A randomly selected player $i$ decides to imitate the strategy of player $j$ who is also randomly selected from all the direct neighbors of player $i$ by comparing their payoff difference with the following probability:

$$W_{i\leftarrow j} = \frac{1}{1+\exp\left(\left(\Pi_i - \Pi_j\right)/K\right)}, \tag{8}$$

where $\Pi_i$ and $\Pi_j$ is the acquired total payoff of the focal player $i$ and its randomly selected neighbor $j$, respectively. $K$ denotes the noise in the imitation process, and we fixed the value of $K$ to be 0.1 throughout the study.

A full Monte Carlo step is to repeat the above procedure $L^2$ times, and $L^2$ is the number of nodes in the given network. Each individual updates their strategy once on average. To subside the transient dynamics and avoid the finite-size effect, we ran simulations for 50000 steps on a regular lattice with sizes ranging from 200×200 to 800×800. The final fraction of each strategy was obtained after up to 45000 steps. The presented data were averaged over 20 independent runs.

## 3 Results

### 3.1 Well-mixed populations

**Finite population.** In finite populations, exiters influence the dynamics between defectors and altruistic punishers, primarily by replacing defectors as the incentive to exit increases. Although altruistic punishers manage to survive, they never dominate; the population is typically controlled by either exiters or defectors. Introducing negative payoffs for exiters transforms the game into a traditional weak prisoner's dilemma, where defectors tend to dominate due to the high cost and low impact of punishment [32, 33]

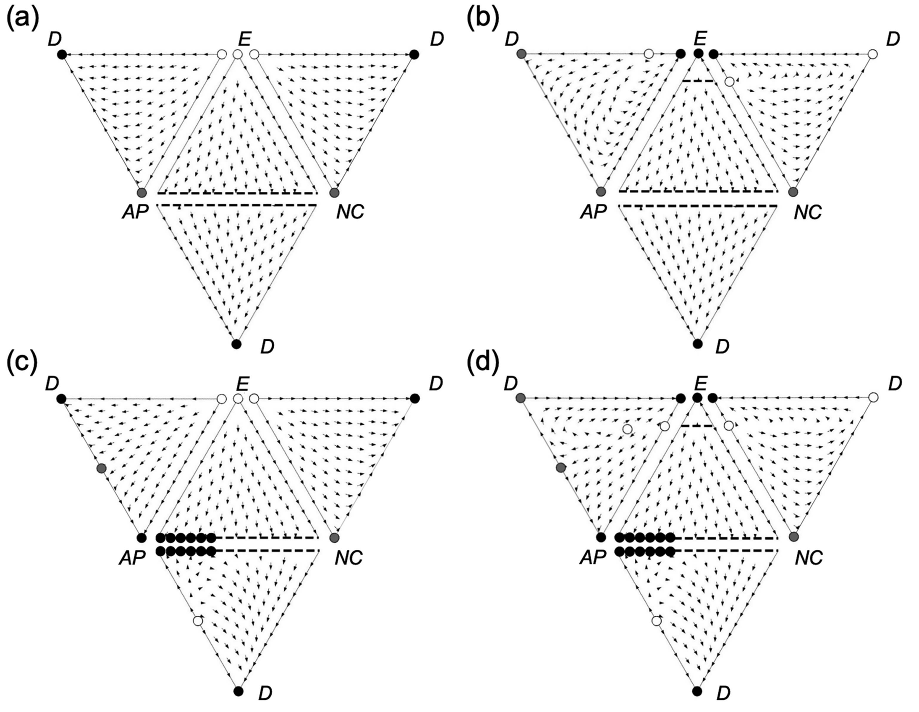

**Figure 2** Adding exit option destabilizes defection regardless of whether altruistic punishment can establish cooperation in an infinite population. Ternary plots along each interface illustrate how adding the exit option destabilizes defection. The dashed line on the AP-NC edge indicates that all the points on this edge are unstable. The filled black circles, filled gray circles, and unfilled circles represent stable fixed points, saddle points, and unstable points, respectively. Parameters: $\beta = 0.3$, $\gamma = 0.1$, (a) $\varepsilon = -0.2$, $b = 1.5$; (b) $\varepsilon = 0.2$, $b = 1.5$; (c) $\varepsilon = -0.2$, $b = 1.2$; (d) $\varepsilon = 0.2$, $b = 1.2$.

(see Figure 1(b)). However, if exit payoffs are slightly positive, a cyclic dominance emerges, allowing altruistic punishers to coexist with defectors and exiters (see Figure 1(c)). This dynamic peaks initially but then declines, leading to the eventual extinction of altruistic punishers (see Figure 1(a)). Simultaneously, this setup supports the existence of non-punishing cooperators through an alternative cyclic dominance route involving non-punishing cooperators, defectors, and exiters, enhancing their survival. Ultimately, while exiters destabilize defectors, their capacity to promote altruistic punishment is limited, indicating that their effectiveness heavily depends on the structure of exit payoffs.

**Infinite population.** Transitioning from finite to infinite populations does not significantly alter the evolutionary outcomes, as evidenced by our stability analysis (theoretical analyses are presented in SI). Below, we present the key findings as theorems to highlight the main results and their implications.

**Theorem 1** (Stability of defection). When the benefit-cost ratio minus the punishment fine $b - \beta$ exceeds 1 and the exit payoff $\epsilon$ is negative, the monomorphic defecting equilibrium is stable. All other equilibria are unstable under these conditions (Figure 2(a)).

**Theorem 2** (Dominance of exiters). When $b - \beta > 1$ and $\epsilon > 0$, the monomorphic exiting equilibrium becomes stable, leading to the dominance of exiters and the instability of other strategies. This result shows how positive exit payoffs can shift the population dynamics towards exiting (Figure 2(b)).

**Theorem 3** (Potential for cooperation). For $b - \beta < 1$ and $\epsilon < 0$, the evolutionary dynamics lead to a mixed equilibrium of altruistic punishers and non-punishing cooperators or revert to the monomorphic defecting equilibrium (Figure 2(c)).

**Theorem 4** (Influence of exit payoffs). If $b - \beta < 1$ and $\epsilon > 0$, the system may stabilize either in a mixed equilibrium involving altruistic punishers and non-punishing cooperators or in a monomorphic exiting equilibrium (Figure 2(d)).

To sum up, Theorems 1 and 2 indicate that exiters cannot support altruistic punishment when punishment alone is insufficient to support cooperation $b - \beta > 1$. In this scenario, defection remains the equilibrium, and incentivizing players to exit only destabilizes defection, replacing it as the equilibrium. Theorems 3 and 4 show that exiters do not hinder the coexistence of altruistic punishment and non-punishing cooperators when punishment can support cooperation $b - \beta < 1$. Incentivizing players to exit

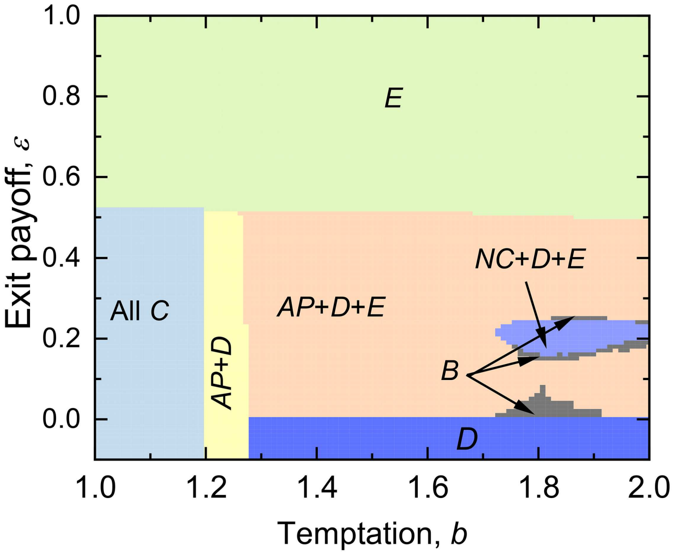

**Figure 3** (Color online) Exit option establishes altruistic punishment in networked population. The full $\epsilon - b$ phase diagram, derived from Monte Carlo simulations of the extended weak prisoner's dilemma game on a regular lattice, is presented. Exiters dominate the population for large exit incentives ($\epsilon \gtrsim 0.51$). Varying incentives lead to six outcomes. For small $b$ values ($b \lesssim 1.19$), altruistic punishment effectively promotes cooperation. Altruistic punishers coexist with defectors for $1.19 \lesssim b \lesssim 1.29$. Larger $b$ values ($b \gtrsim 1.29$) result in full defection with negative $\epsilon$, while positive $\epsilon$ leads to the coexistence of altruistic punishers with defectors and exiters, second-order free riders with defectors and exiters, or a bi-stable state between these coexistence types.

destabilizes defection, allowing it to coexist with altruistic punishment and non-punishing cooperators.

Having established the impact of exiters on altruistic punishment in both finite and infinite populations, we now turn our attention to networked populations. The structure and connectivity of networks introduce network reciprocity [34], where cooperators form compact clusters to resist the invasion of defectors. This leads us to expect that social networks can potentially alter the dynamics of exiters and their impact on cooperation and altruistic punishment. By examining these interactions within networked populations, we aim to determine whether exiters can explain and promote altruistic punishment under structured populations.

### 3.2 Networked population

Figure 3 shows the full $\epsilon - b$ phase diagram obtained by extensive Monte Carlo simulations. The addition of the simple exit option leads to complicated evolutionary outcomes. Initially, when the incentives to exiters are sufficiently large, $\epsilon \gtrsim 0.51$, the exiters outcompete other action types and dominate the whole population (the E phase in Figure 3), consistent with previous findings in the absence of altruistic punishers [24]. With smaller incentives to exiters, $\epsilon \lesssim 0.51$, leads to six different possible outcomes. If the temptation to defect is relatively small, $b \lesssim 1.29$, altruistic punishment together with network reciprocity is sufficient to maintain prosocial behavior (the All C phase and the AP + D phase in Figure 3).

When $b \lesssim 1.19$, defectors can be completely eliminated by altruistic punishers, allowing altruistic punishers and non-punishing cooperators to coexist in a regular lattice. In the absence of defectors, non-punishing cooperators and altruistic punishers are indistinguishable, and the evolutionary dynamics may lead to the full AP state, the full NC state, or the mixed AP + NC state, determined by initial conditions (the All C phase in Figure 3). With increasing $b$, $1.19 \lesssim b \lesssim 1.29$, the effectiveness of altruistic punishment is reduced, and defectors coexist with altruistic punishers (the AP + D phase in Figure 3).

If $b$ is sufficiently large, altruistic punishment loses its effectiveness in sustaining cooperation alongside network reciprocity, and defectors dominate the entire population for negative $\epsilon$ (the D phase in Figure 3). Non-negative values of exit payoffs convert the full defection state to three possible outcomes: (i) the coexistence of AP, D, and E (the AP + D + E phase in Figure 3), (ii) the coexistence of NC, D, and E (the NC + D + E phase in Figure 3), or (iii) a bi-stable state between these two coexistences (the B phase in Figure 3).

In summary, in structured populations, when altruistic punishment alone can outcompete defectors and favor cooperation on a square lattice, the introduction of an exit option does not affect the coexistence of

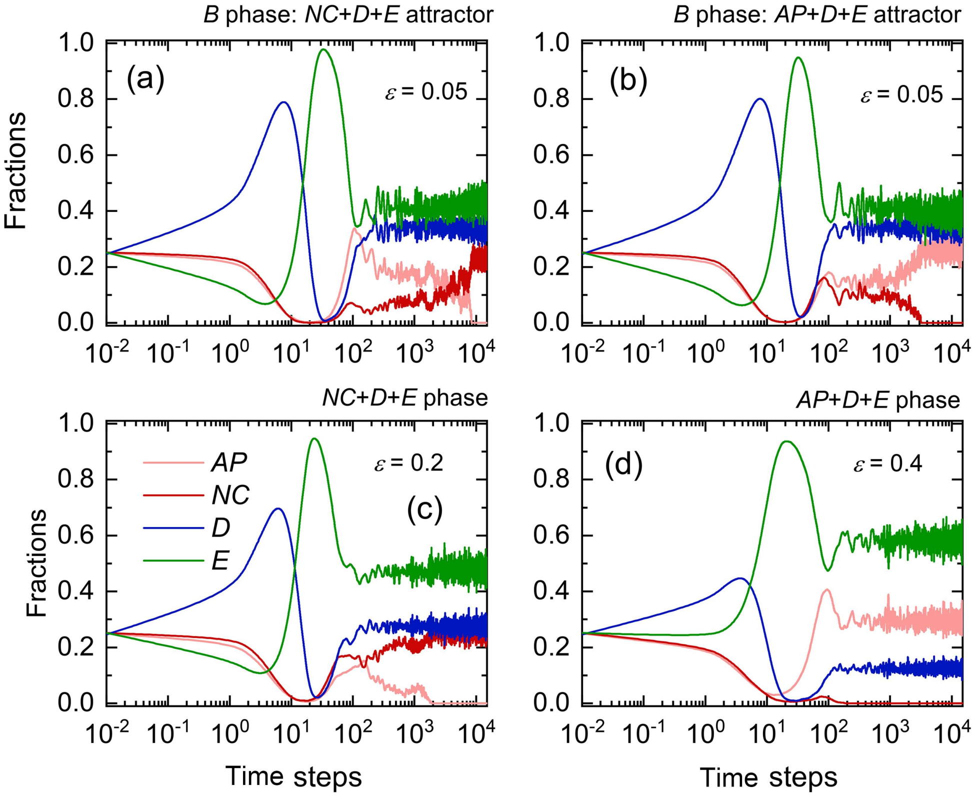

**Figure 4** (Color online) Time dependence of actor abundances reveals complicated evolutionary dynamics. In the bi-stable phase, starting from random initial conditions, small incentives for exit options lead the system to either the NC + D + E or AP + D + E attractor, but coexistence of all four actors is not feasible. Depending on the initial abundance ratio of altruistic punishers to non-punishing cooperators, either altruistic punishers or non-punishing cooperators coexist with defectors and exiters through cyclic dominance ((a) or (b)). Larger incentives to exiters transition the bi-stability to monostability, with evolutionary outcomes determined by exit incentives. Parameters: (a) and (b) $b = 1.8$, $\epsilon = 0.05$, (c) $\epsilon = 0.2$, and (d) $\epsilon = 0.4$.

altruistic punishers and non-punishing cooperators. However, when altruistic punishment alone cannot outcompete defectors and defection dominates, exiters play a crucial role in fostering altruistic punishment within the networked population. The effectiveness of this role is sensitive to exit payoffs, which can lead to the replacement of altruistic punishment by non-punishing cooperators or result in a bi-stable state between these coexistence types.

To gain a better understanding of how these actors coexist in the population, the evolution features of the fractions of each actors was examined and the results are presented in Figure 4. In the bi-stable phase, it is the cooperators (altruistic punishers or non-punishing cooperators) start giving way to the defectors and with fewer cooperators around, defectors then giving way to the exiters. With large numbers of exiters, both the altruistic punishers and non-punishing cooperators compete for the exiters as they can only survive by adhering to the exiters. The described phenomenon is the cyclic dominance in which these actors dominate one another.

There are two possible routes of cyclic dominance: (i) altruistic punishers dominate exiters, who dominate defectors, who in turn dominate altruistic punishers; or (ii) non-punishing cooperators dominate exiters, who dominate defectors, who then dominate non-punishing cooperators. Researchers have confirmed the effectiveness of cyclic dominance in maintaining biodiversity or fostering cooperation [35, 36].

Despite starting with random initial conditions, the evolutionary outcomes vary when conducting multiple independent simulations under the same parameter settings. For example, in the NC + D + E attractor (Figure 4(a)), the fraction of altruistic punishers is temporarily much larger than that of non-punishing cooperators at around the 100th step, then the faction of altruistic punishers gradually decreases until it is eliminated and the fraction of second-order free riders increases until it reaches a stable state. However, in the AP + D + E attractor (Figure 4(b)), the fraction of altruistic punishers is always comparable to that of non-punishing cooperators up to around the 1000th step. After this critical time step, the fraction of non-punishing cooperators gradually decreases until it is eliminated, and altruistic

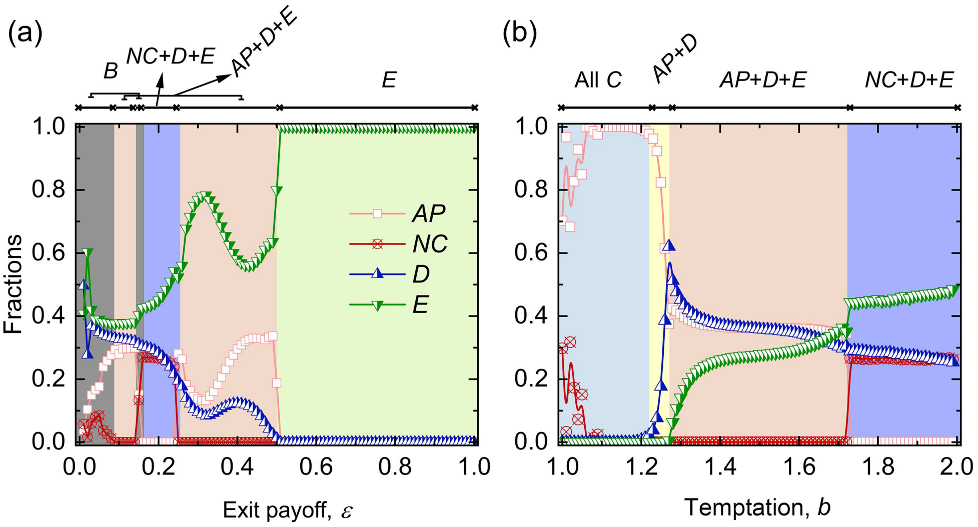

**Figure 5** (Color online) Power relations between altruistic punishers, second-order free riders, defectors, and exiters exhibiting complicated equilibria. (a) Along the vertical transect of $\epsilon - b$ phase at $b = 1.8$; (b) along the horizontal transect of $\epsilon - b$ phase plane at $\epsilon = 0.2$.

punishers gradually increase to reach a stable state. Thus, the indirect competition, possibly influenced by their ability to adhere to exiters (which may be related to their densities in the population), determines the fate of altruistic punishers and non-punishing cooperators. We will investigate this aspect shortly.

The phenomenon of bistability disappears when the incentives for exiters are increased. The evolutionary dynamics result in two distinct phases: the NC + D + E phase and the AP + D + E phase, depending on the incentives for exiters. In both cases, altruistic punishers and non-punishing cooperators can dominate over exiters when the fraction of exiters peaks. However, when the incentives for exiters are at an intermediate level ($\epsilon = 0.2$), it is the non-punishing cooperators that dominate over exiters. Altruistic punishers lose in indirect competition with non-punishing cooperators and are eventually eliminated as the simulation proceeds. In this scenario, non-punishing cooperators coexist with defectors and exiters through cyclic dominance in the networked population (Figure 4(c)). On the other hand, if the incentives for exiters are higher ($\epsilon = 0.4$), altruistic punishers begin to dominate over exiters when exiters reach their first peak, and the non-punishing cooperators are unable to surpass the exiters and eventually get eliminated. In this case, altruistic punishers coexist with defectors and exiters through cyclic dominance in the system (Figure 4(d)).

To analyze how exiters affect the quantitative power relationships among different actors, we present two representative cross sections of the phase diagram in Figure 5. Figure 5(a) displays the stationary fractions of the four competing actors along the vertical transect of the $\epsilon - b$ phase plane, with $b = 1.8$. When the exit is costly ($\epsilon < 0$), the defectors dominate the whole population (the D phase in Figure 1). As shown in Figure 5(a), if the incentives to exiters are small but positive, the D phase gives way to the B phase, where the system converges to either the AP + D + E attractor or the NC + D + E attractor depending on the results of the indirect competition between the altruistic punishers and non-punishing cooperators. By further increasing the $\epsilon$, the NC + D + E phase is reached at $\epsilon \approx 0.17$, and there are two narrow strips that the AP + D + E phase and B phase can dominate separately during this increment. The AP + D + E phase dominates in the range $0.06 \lesssim \epsilon \lesssim 0.16$, and the B phase is short-lived again in the range $0.16 \lesssim \epsilon \lesssim 0.17$. As $\epsilon$ continues to increase, the NC + D + E phase gives way to AP + D + E phase via discontinuous phase transition at $\epsilon \approx 0.25$. When incentives to exiters are sufficiently large, the AP + D + E phase is finally replaced by the E phase at the critical point $\epsilon \approx 0.51$.

Figure 5(b) shows the horizontal transect of $\epsilon - b$ at $\epsilon = 0.2$, it also reveals the power relations between these competing actors, but it is dependent on the temptation level, $b$. When $b$ is small, $1 \leqslant b \lesssim 1.29$, the altruistic punishment together with the network reciprocity is able to support prosocial behavior. When $1 \leqslant b \lesssim 1.23$, the altruistic punishers can completely eliminate the defectors, the elimination of the defectors also negatively affects the exiters, and thus altruistic punishers coexist with non-punishing cooperators as they cannot be distinguished in the absence of defectors. The All C phase transitions to the AP + D phase through a continuous phase transition. Although the advantages of cooperators

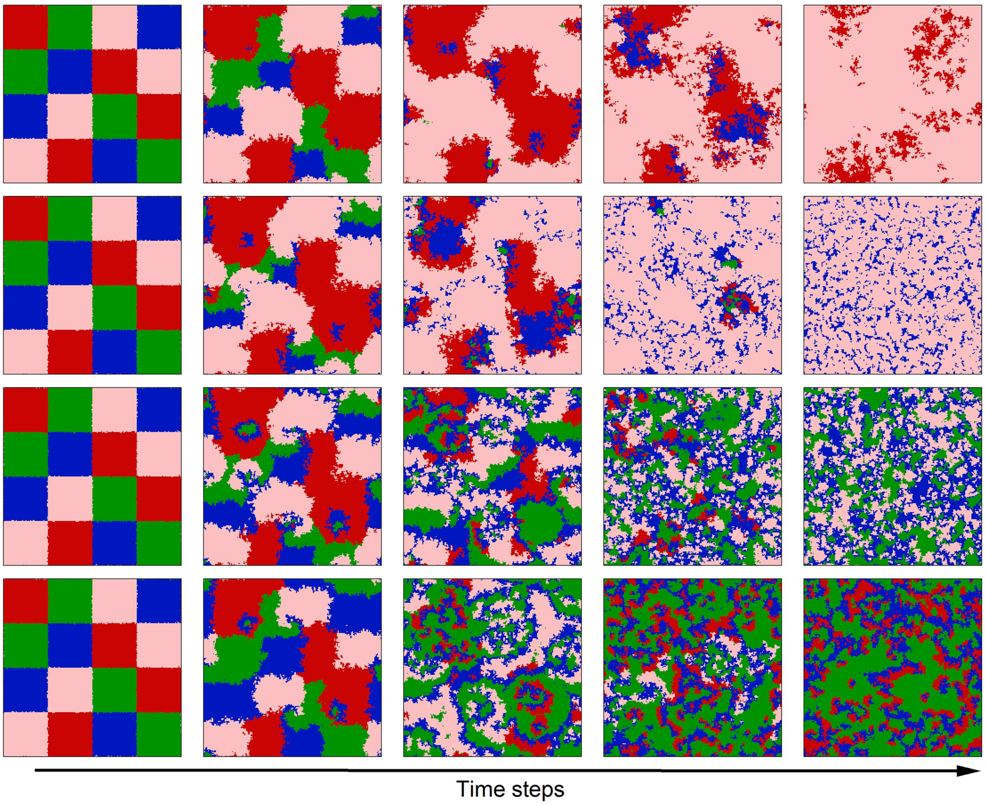

**Figure 6** (Color online) Evolutionary snapshots illustrating dominance modes between actors. Shown are snapshots at different time steps (columns) and for varying temptations for defection (rows). The snapshots depict the dynamics where exiters, defectors, altruistic punishers, and non-punishing cooperators interact. Results were obtained with $\epsilon = 0.2$ after the 30000th step to generate the final snapshots (rightmost column). The intermediate snapshots (the first to fourth columns) were taken at different time steps across rows for illustration purposes. From top to bottom, the values of $b$ are fixed at 1.04, 1.25, 1.5, and 1.8, respectively.

decrease with increasing $b$, those cooperators who punish defectors gain a greater advantage compared to defectors. Therefore, in this scenario, network reciprocity supports the coexistence of altruistic punishers and defectors. If the conditions to support cooperation with altruistic punishment are unfavorable, adding an exit option can promote the system to the AP + D + E phase when $b \lesssim 1.73$. However, with increasing $b$, the AP + D + E phase gives way to the NC + D + E phase through a discontinuous phase transition at the critical point, $b \approx 1.73$.

To thoroughly examine the evolutionary dynamics and the indirect competition between altruistic punishers and non-punishing cooperators in both spatial and temporal dimensions, we plotted evolutionary snapshots for different values of $b$ at $\epsilon = 0.2$, as shown in Figure 6.

When the temptation is low (the top row in Figure 6), exiters are initially eliminated by both altruistic punishers and non-punishing cooperators. Shortly after, defectors face the same fate. Eventually, altruistic punishers and non-punishing cooperators coexist as they become indistinguishable, leading to a frozen state in the system. As temptation increases (the second row in Figure 6), defectors become more competitive. They can coexist with altruistic punishers instead of being eliminated. However, this coexistence does not ensure the survival of exiters, who are still eliminated when temptation is low. Defectors eventually eliminate non-punishing cooperators, leading to the coexistence of altruistic punishers and defectors. With even higher temptation (the third row in Figure 6), defectors can invade both altruistic punishers and non-punishing cooperators. They negatively impact non-punishing cooperators more than altruistic punishers, leading to the elimination of non-punishing cooperators first. This results in the coexistence of altruistic punishers, defectors, and exiters. At the highest temptation level (the bottom row in Figure 6), defectors are the most competitive. Altruistic punishers and non-punishing cooperators are exploited by defectors at a similar pace, leading to the dominance of exiters by eliminating defectors. In the indirect competition between exiters and non-punishing cooperators, altruistic punishers lose their advantages due to the cost of punishment. Consequently, non-punishing cooperators coexist with defectors and exiters.

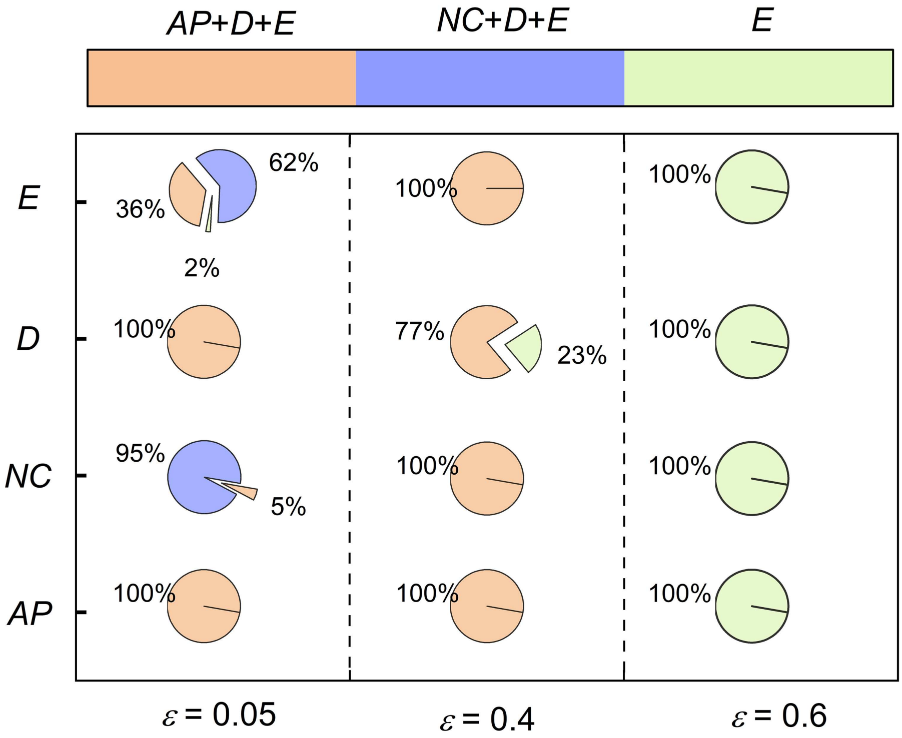

**Figure 7** (Color online) Initial conditions determine the outcome of altruistic punishers and non-punishing cooperators in the bi-stable phase. Shown are the evolutionary outcomes after implementing 100 independent simulations for each parameter combination under four different initial conditions. The initial conditions were (i) 97% of players were initially assigned as AP, (ii) 97% of players were initially assigned as NC, (iii) 97% of players were initially assigned as D, and (iv) 97% of players were initially assigned as E. The rest of the other action types were assigned to the other players with equal probability in these different initial conditions. Parameters were fixed as $b = 1.8$, from left to right, $\epsilon = 0.05, 0.4, 0.6$, respectively.

Overall, both altruistic punishers and non-punishing cooperators can coexist with defectors and exiters through different cyclic dominance routes. However, these two types of cyclic dominance cannot coexist simultaneously in the population. The indirect competition for the territories of exiters between altruistic punishers and non-punishing cooperators determines the outcome of the competition. Competitive defectors more easily impact non-punishing cooperators negatively than altruistic punishers, leading to their elimination first in certain scenarios.

It is generally accepted that initial conditions, such as different initial densities of each actor, are crucial for evolutionary outcomes in agent-based models [37]. We further assessed how the initial fractions of actors shape the evolutionary outcomes in the bi-stability phase. Figure 7 presents the evolutionary outcomes for $\epsilon = 0.05$, 0.4, and 0.6 under four different initial conditions: (i) 97% of players were initially assigned as AP (altruistic player), (ii) 97% of players were initially assigned as NC (non-cooperative), (iii) 97% of players were initially assigned as D (defector), and (iv) 97% of players were initially assigned as E (exiter). The remaining players were assigned one of the other three actions with equal probability in these conditions. The results were obtained by implementing 100 independent simulations. At $\epsilon = 0.05$ (the left column in Figure 7), the system consistently settled into the AP + D + E attractor if the majority initially were AP or D. However, with a majority of NC initially, the system favored the NC + D + E attractor with a 95% probability. The initial majority as E led to a 36% probability of the AP + D + E and 62% probability of the NC + D + E attractor. Larger exit incentives shifted from bi-stability to monostability (the middle and right columns in Figure 7), where outcomes were either AP + D + E or E depending solely on exit incentives, irrespective of initial conditions. Notably, at $\epsilon = 0.4$, there was a 23% probability of the system falling into the E phase when most players were initially D (the middle column in Figure 7). This unexpected result might be due to finite-size effects, which can mislead results in structured populations [37]. Therefore, using a larger network or subsystem solutions is crucial to avoid this issue [38, 39]. We believe the counterintuitive E phase is due to finite-size effects and expect a pure AP + D + E phase with a larger network size.

## 4 Discussion

To conclude, in this paper, we have shown that exiters lead to complicated dynamics for altruistic punishers. In well-mixed populations (whether finite or infinite), exiters destabilize defection and gradually replace defectors as the payoff to exit increases, but they fail to encourage the prevalence of altruistic punishment. This effect is consistent with their role in cooperation when the altruistic punishment option is absent in well-mixed scenarios. However, in social networks with non-negative intermediate incentives for exiters (i.e., less than 0.5 in our results), exiters generally facilitate the coexistence of altruistic punishment through a cyclic dominance effect involving altruistic punishers, defectors, and exiters. This cyclic dominance is unstable and sensitive to the incentives for exiters. Adjustments to the exit payoff can lead to new cyclic dominance where non-punishing cooperators replace altruistic punishers or to bi-stability between these two cycles. In the bi-stable outcome, the system converges to the coexistence of altruistic punishers if defectors or altruistic punishers initially dominate the population. Otherwise, the outcomes are more likely to converge to the coexistence of non-punishing cooperators if non-punishing cooperators or exiters initially dominate the population. These results reveal the nuanced effect of exiters in encouraging altruistic punishers, highlighting the need for carefully incentivizing exit behavior to support altruistic punishment, as their positive impact is sensitive to population structure and exit payoffs.

Our results have multiple interesting implications. From a broader socioeconomic perspective, the findings suggest that the strategic use of exit options can intricately influence cooperative behaviors and the sustainability of altruistic punishment. Policy makers could consider implementing exit options in community programs to disrupt harmful behaviors and foster cooperation. For example, in organizational management, companies could design work environments where employees have the flexibility to exit or switch tasks if they encounter uncooperative peers, promoting a healthier and more productive workplace culture. However, our results also highlight potential negative impacts: exiters can sometimes replace altruistic punishers with non-punishing cooperators, which may weaken the enforcement of cooperative norms. This indicates that while exit options can promote cooperation, they can also inadvertently undermine the mechanisms that sustain it, depending on the context and incentives. Therefore, recognizing the sensitivity of exit behavior to population structure and incentives is crucial. Strategies must be carefully crafted to ensure that exit options enhance rather than detract from cooperative behavior, balancing the benefits of flexibility with the need to maintain strong cooperative norms.

Parallels between our exiters and well-known loners, both forms of voluntary participation, invite comparisons regarding their role in altruistic punishment. However, interpreting our results in that context is challenging due to the lack of a direct comparison with loners. The positive role of loners on altruistic punishers has primarily been studied in public goods games with multiplayer interactions, while exiters were initially focused on two-player games. Nonetheless, we can draw some qualitative conclusion: loners have a well-documented positive influence on altruistic punishers in well-mixed populations [21, 23], whereas exiters exhibit a more nuanced effect that is sensitive to population structure and exit payoffs. These similarities and differences confirm the positive role of voluntary participation in promoting altruistic punishment but also suggest that loners are more effective than exiters.

Returning to a narrower perspective on the impact of voluntary participation on altruistic punishment, for the results to be truly useful in policy making, the role of voluntary participation in altruistic punishment needs to be further revisited by examining this mechanism itself. The central idea of voluntary participation is that the game is not compulsory, and players can choose to forgo the benefits from public goods and exit the game. However, how such exit behavior influences public goods or defines the payoff of such behavior is quite flexible. Both loners and exiters are specific forms of voluntary participation, and other forms should also be expected. As voluntary participation is a bottom-up scheme, its potential impact could also be related to exiters' personalities. For example, according to the concept of social value orientation [40], prosocial exiters could bring additional benefits to public goods (reflecting their concern for others), individualistic exiters could focus solely on their own payoff without benefiting public goods (similar to loners), and competitive exiters could damage public goods (reflecting their desire to gain more than others, somewhat analogous to our exiters). Beyond these, the social value orientation framework suggests additional motivations for innovative variants of voluntary strategies. These include masochism, where individuals accept negative payoffs by exiting the game without affecting others; martyrdom, which entails negative personal payoffs alongside generating positive outcomes for others; and sadomasochism, characterized by negative personal payoffs coupled with inflicting harm on opponents, among others. Therefore, developing a comprehensive theoretical model that refines voluntary participation strategies by considering personality presents a compelling research direction for understanding the availability and effectiveness of voluntary participation in altruistic punishment, particularly given the heterogeneous impact of exiters on public goods.

Future investigations should also examine the robustness of our findings across diverse network structures, considering the full set of punishment mechanisms, as well as conducting human behavior experiments.

Our results were primarily obtained on a square lattice, which, although simple and capable of capturing the basic features of social networks, does not fully represent the characteristics of real social networks that exhibit heterogeneity [41], higher-order connections [42], or temporal dynamics [43]. Validating our results on more realistic social networks would enhance our understanding of the impact of exiters on altruistic punishment. Second, our model only considers non-punishing cooperators, who are one of the factors challenging the stability of altruistic punishment. In reality, the effectiveness of altruistic punishment is also challenged by antisocial punishment, which has been observed in various human cultures through experimental studies [17,44,45]. Additionally, if punishment could target loners or if loners could punish other actors, it would provide further insights into altruistic punishment [46]. Therefore, extending our model to include the full set of punishment interactions among actors would provide additional insights.

Finally, it is crucial to further verify our findings through human behavior experiments. Experimental studies often produce contrasting or surprising results compared to theoretical predictions. For instance, a recent study explored the introduction of punishment into networks as a means to theoretically promote cooperation [33,47–49]. However, a large-scale human behavior experiment concluded that the introduction of peer punishment did not promote cooperation in structured populations but instead reduced the benefits of network reciprocity [50]. While our study supports the idea that exiters can facilitate the prevalence of altruistic punishment when integrated into networks, it remains imperative to design experimental models that can further validate our theory through human behavior experiments.

**Acknowledgements** This work was supported by National Key R&D Program of China (Grant No. 2022YFE0112300) and National Science Fund for Distinguished Young Scholars (Grant No. 62025602). We also acknowledge support from (i) JSPS Postdoctoral Fellowship Program for Foreign Researchers (Grant No. P21374), Accompanying Grant-in-aid for Scientific Research from JSPS KAKENHI (Grant No. JP 22F31374), and National Natural Science Foundation of China (Grant No. 11931015) to Chen SHEN as a co-investigator, (ii) National Natural Science Foundation of China (Grant No. 11931015) and Major Program of National Fund of Philosophy and Social Science of China (Grants Nos. 22&ZD158, 22VRCO49) to Lei SHI, (iii) National Natural Science Foundation of China (Grant Nos. U22B2036, 11931015), Key Technology Research and Development Program of Science and Technology-Scientific and Technological Innovation Team of Shaanxi Province (Grant No. 2020TD-013), and XPLORER PRIZE. to Zhen WANG, and (iv) Grant-in-aid for Scientific Research from JSPS KAKENHI (Grants Nos. JP 20H02314, JP 23K28189, JP 23H03499) awarded to Jun TANIMOTO. We thank Prof. Dr. Marko JUSUP for valuable discussions.

**Supporting information** Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

### References

1 Grossmann I, Feinberg M, Parker D C, et al. AI and the transformation of social science research. Science, 2023, 380: 1108–1109

2 Crandall J W, Oudah M, Tennom M, et al. Cooperating with machines. Nat Commun, 2018, 9: 233

3 Wang L, Fu F, Chen X R. Mathematics of multi-agent learning systems at the interface of game theory and artificial intelligence. Sci China Inf Sci, 2024, 67: 166201

4 Hardin G. The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality. Science, 1968, 162: 1243–1248

5 Nowak M A. Five rules for the evolution of cooperation. Science, 2006, 314: 1560–1563

6 Rand D G, Nowak M A. Human cooperation. Trends Cogn Sci, 2013, 17: 413–425

7 Fehr E, Gächter S. Altruistic punishment in humans. Nature, 2002, 415: 137–140

8 Fehr E, Gächter S. Cooperation and punishment in public goods experiments. Am Economic Rev, 2000, 90: 980–994

9 Andreoni J, Harbaugh W, Vesterlund L. The carrot or the stick: rewards, punishments, and cooperation. Am Economic Rev, 2003, 93: 893–902

10 Hilbe C, Sigmund K. Incentives and opportunism: from the carrot to the stick. Proc R Soc B, 2010, 277: 2427–2433

11 Wang Z, Jusup M, Shi L, et al. Exploiting a cognitive bias promotes cooperation in social dilemma experiments. Nat Commun, 2018, 9: 2954

12 Hauert C, de Monte S, Hofbauer J, et al. Replicator dynamics for optional public good games. J Theor Biol, 2002, 218: 187–194

13 Hauert C, de Monte S, Hofbauer J, et al. Volunteering as red queen mechanism for cooperation in public goods games. Science, 2002, 296: 1129–1132

14 Sasaki T, Uchida S. The evolution of cooperation by social exclusion. Proc R Soc B, 2013, 280: 20122498

15 Han T A, Pereira L M, Lenaerts T. Evolution of commitment and level of participation in public goods games. Auton Agent Multi-Agent Syst, 2017, 31: 561–583

16 Shen C, He Z, Shi L, et al. Prosocial punishment bots breed social punishment in human players. J R Soc Interface, 2024, 21: 20240019

17 Herrmann B, Thöni C, Gächter S. Antisocial punishment across societies. Science, 2008, 319: 1362–1367

18 Rand D G, Armao IV J J, Nakamaru M, et al. Anti-social punishment can prevent the co-evolution of punishment and cooperation. J Theor Biol, 2010, 265: 624–632

19 Chaudhuri A. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Exp Econ, 2011, 14: 47–83

20 Wang L, Cong R, Li K. Feedback mechanism in cooperation evolving (in Chinese). Sci Sin Inform, 2014, 44: 1495–1514

21 Brandt H, Hauert C, Sigmund K. Punishing and abstaining for public goods. Proc Natl Acad Sci USA, 2006, 103: 495–497

22 Fowler J H. Altruistic punishment and the origin of cooperation. Proc Natl Acad Sci USA, 2005, 102: 7047–7049

23 Hauert C, Traulsen A, Brandt H, et al. Via freedom to coercion: the emergence of costly punishment. Science, 2007, 316: 1905–1907

24 Shen C, Jusup M, Shi L, et al. Exit rights open complex pathways to cooperation. J R Soc Interface, 2021, 18: 20200777

25 Nowak M A. Evolutionary Dynamics: Exploring the Equations of Life. Cambridge: Harvard University Press, 2006

26 Fudenberg D, Imhof L A. Imitation processes with small mutations. J Economic Theor, 2006, 131: 251–262

27 Imhof L A, Fudenberg D, Nowak M A. Evolutionary cycles of cooperation and defection. Proc Natl Acad Sci USA, 2005, 102: 10797–10800

28 Santos F C, Pacheco J M, Skyrms B. Co-evolution of pre-play signaling and cooperation. J Theor Biol, 2011, 274: 30–35

29 Nowak M A, Sasaki A, Taylor C, et al. Emergence of cooperation and evolutionary stability in finite populations. Nature, 2004, 428: 646–650

30 Traulsen A, Nowak M A, Pacheco J M. Stochastic dynamics of invasion and fixation. Phys Rev E, 2006, 74: 011909

31 Sigmund K, de Silva H, Traulsen A, et al. Social learning promotes institutions for governing the commons. Nature, 2010, 466: 861–863

32 Egas M, Riedl A. The economics of altruistic punishment and the maintenance of cooperation. Proc R Soc B, 2008, 275: 871–878

33 Szolnoki A, Perc M. Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment. Phys Rev X, 2017, 7: 041027

34 Nowak M A, May R M. Evolutionary games and spatial chaos. Nature, 1992, 359: 826–829

35 Reichenbach T, Mobilia M, Frey E. Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games. Nature, 2007, 448: 1046–1049

36 Szolnoki A, Mobilia M, Jiang L L, et al. Cyclic dominance in evolutionary games: a review. J R Soc Interface, 2014, 11: 20140735

37 Perc M, Jordan J J, Rand D G, et al. Statistical physics of human cooperation. Phys Rep, 2017, 687: 1–51

38 Szolnoki A, Perc M. Competition of tolerant strategies in the spatial public goods game. New J Phys, 2016, 18: 083021

39 Perc M. Stability of subsystem solutions in agent-based models. Eur J Phys, 2017, 39: 014001

40 Murphy R O, Ackermann K A, Handgraaf M J J. Measuring social value orientation. Judgm decis Mak, 2011, 6: 771–781

41 Santos F C, Santos M D, Pacheco J M. Social diversity promotes the emergence of cooperation in public goods games. Nature, 2008, 454: 213–216

42 Xu Y, Wang J, Xia C Y, et al. Higher-order temporal interactions promote the cooperation in the multiplayer snowdrift game. Sci China Inf Sci, 2023, 66: 222208

43 Li A, Zhou L, Su Q, et al. Evolution of cooperation on temporal networks. Nat Commun, 2020, 11: 2259

44 Denant-Boemont L, Masclet D, Noussair C N. Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. Economic Theor, 2007, 33: 145–167

45 Gächter S, Herrmann B. The limits of self-governance when cooperators get punished: experimental evidence from urban and rural Russia. Eur Economic Rev, 2011, 55: 193–210

46 Rand D G, Nowak M A. The evolution of antisocial punishment in optional public goods games. Nat Commun, 2011, 2: 434

47 Brandt H, Hauert C, Sigmund K. Punishment and reputation in spatial public goods games. Proc R Soc Lond B, 2003, 270: 1099–1104

48 Helbing D, Szolnoki A, Perc M, et al. Punish, but not too hard: how costly punishment spreads in the spatial public goods game. New J Phys, 2010, 12: 083005

49 Perc M, Szolnoki A. Self-organization of punishment in structured populations. New J Phys, 2012, 14: 043013

50 Li X, Jusup M, Wang Z, et al. Punishment diminishes the benefits of network reciprocity in social dilemma experiments. Proc Natl Acad Sci USA, 2018, 115: 30–35