PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

# Bilateral costly expulsions resolve the public goods dilemma

Xiaofeng Wang[1,2] and Matjaž Perc[3,4,5,6]

[1]Department of Automation, School of Information Science and Technology, Donghua University, Shanghai 201620, China
[2]Engineering Research Center of Digitized Textile and Apparel Technology, Donghua University, Ministry of Education, Shanghai 201620, China
[3]Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, Maribor 2000, Slovenia
[4]Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404332, Taiwan
[5]Alma Mater Europaea ECM, Slovenska ulica 17, Maribor 2000, Slovenia
[6]Complexity Science Hub Vienna, Josefstädterstraße 39, Vienna 1080, Austria

MP, 0000-0002-3087-541X

Research

**Cite this article:** Wang X, Perc M. 2021 Bilateral costly expulsions resolve the public goods dilemma. *Proc. R. Soc. A* **477**: 20210627.
https://doi.org/10.1098/rspa.2021.0627

Received: 2 August 2021
Accepted: 24 August 2021

**Subject Areas:**
complexity

**Keywords:**
social network, evolutionary game theory, cooperation, expulsion, public goods, complexity

**Author for correspondence:**
Matjaž Perc
e-mail: matjaz.perc@gmail.com

Expulsion has been found to promote cooperation in social dilemmas, but only if it does not incur costs or is applied unilaterally. Here, we show that removing both conditions leads to a spontaneous resolution of the costly expulsion problem. Namely, by studying the public goods game where cooperators and defectors can expel others at a personal cost, we find that public cooperation thrives as expulsion costs increase. This is counterintuitive, as the cost of other-regarding behaviour typically places an additional burden on cooperation, which is in itself costly. Such scenarios are referred to as second-order free-rider problems, and they typically require an additional mechanism, such as network reciprocity, to be resolved. We perform a mean field analysis of the public goods game with bilateral costly expulsion, showing analytically that the expected payoff difference between cooperators and defectors increases with expulsion costs as long as players with the same strategy have, on average, a higher frequency to interact with each other. As the latter condition is often satisfied in social networks, our results thus reveal a fascinating new path to public cooperation, and they show that the costs of well-intended actions need not be low for them to be effective.

## 1. Introduction

Social dilemma games, wherein each player can adopt either cooperation or defection to interact with its opponents, are general metaphors for studying the evolution of cooperation under the framework of evolutionary game theory [1]. Traditionally, the evolutionary dynamics of cooperation is investigated in the context of two-person social dilemma games that are governed by pairwise interactions [2–9], as dictated by complex networks [10,11]. However, the interaction patterns of agents in most, if not all, models of social, economics and biological systems usually fall into the category of group interactions which, in general, cannot be reduced to the corresponding sum of pairwise interactions [12]. Therefore, the public goods game is employed by researchers to study the evolutionary dynamics of public cooperation that is beyond pairwise interactions [13]. In the $n$-person public goods game, each player decides simultaneously whether to invest a certain amount $c > 0$ into a common pool. Then, the total amount of contributions will be multiplied by an enhancement factor (i.e. the marginal return in the classical game theory) $r \in (1, n)$, and the resulting public goods will be equally divided among all $n$ members of the game irrespective of whether they contribute or not. Because a contributor obtains negative net returns for the investment it makes (i.e. $(r/n - 1)c < 0$), rational players should refuse to make any investments so as to maximize their short-term benefits. Therefore, no public goods will be produced, which in turn leads to no collective benefits distributed to any player. However, had each player contributed, everyone would have profited and received a net benefit of $(r - 1)c > 0$. Obviously, the public goods game characterizes well a typical social dilemma—the conflict of interests between the individual and the group [14].

Over the last couple of decades, numerous mechanisms have been discovered by scientists to resolve the public goods dilemma [15,16]. In particular, recent studies have shown that population structure described by networks has profound effects on the evolutionary dynamics of public cooperation. For instance, Santos *et al.* have shown that network structure, including regular and scale-free graphs, is able to promote the evolution of cooperation in the public goods game [17]. Moreover, due to their simple topology structures, the spatial networks are often employed by statistical physicists as important benchmarks to study the spatiotemporal dynamics of observed solutions, the formation of spatial patterns, and self-organization processes that affect the population states of cooperation [18]. For example, Szolnoki *et al.* have found that the indirect linkage of individuals induced by group interactions leads to the disappearance of optimal conditions for the survival of cooperators that do exist for pairwise interactions in the spatial public goods game [19]. Meanwhile, impacts of coevolutionary rules [20], such as migration [21,22] and aspiration [23], in the evolution of cooperation have also been extensively explored in the context of the spatial public goods game. In particular, Wang *et al.* have recently discovered that expulsion costs undermine the resistance ability of positive assortment among cooperative individuals against mutations, though robust cooperation can indeed be realized by assigning ability of expulsion exclusively to cooperators in the spatial public goods game [24].

In this article, we will study the effects of bilateral costly expulsion by endowing both cooperators and defectors with expulsive ability and mainly focus on investigating how expulsion costs affect the evolutionary dynamics of cooperation in the spatial public goods game with bilateral costly expulsion. Herein, we would like to point out that it is necessary for us to introduce costly expulsion in a bilateral manner. Because only in this way can we impartially evaluate the effects of costly expulsion on the evolution of public cooperation. Furthermore, defectors, as rational players, do have the incentive to adopt expulsive behaviours towards other defectors in the public goods game. Because, for each defector in a particular group, the expulsion towards other defectors can lead to the enhancement of the proportion of contributors in this group. Interestingly, as we will show, while bilateral costless expulsion is still positively related to cooperation, expulsion costs can further promote the evolution of public cooperation if applied bilaterally.

## 2. Model

The proposed model of evolutionary games governed by group interactions is constructed on a square lattice of size $L \times L$ with periodic boundary conditions as well as von Neumann neighbourhood (i.e. the node degree $k=4$). Each node of the spatial network represents a site that may or may not be occupied by an individual. This means that the population size $N$ is no more than the number of sites $L^2$ on the square lattice, which leads to the population density $\rho = N/L^2 \leq 1$.

At the start of evolution, a number of $N$ individuals are randomly distributed to $N$ sites of the square lattice, and each individual on a site $i$ is designated as either an expulsive cooperator ($S_i = \mathrm{EC}$) or an expulsive defector ($S_i = \mathrm{ED}$) with equal probability. After the application of random initial conditions, we use the method of Monte Carlo simulation with each time step comprising the following three elementary procedures. First, each player on site $i$ synchronously accumulates its payoff $P_i$ by playing the public goods games centred on both its neighbouring players and itself, if there is indeed any player in the focal player's neighbouring sites. Otherwise, this player obtains nothing. Next, each defector is selected once to be expelled to any empty site on the spatial network in a random sequence manner, provided that at least one expulsive individual has ever played the public goods game with it and that not all sites are occupied by individuals. Here, expulsion refers to the behaviour that is used by both expulsive cooperators and expulsive defectors to expel their free-riding group members from present sites to any other ones on a graph. Note that not only cooperators but also defectors in the public goods game have incentives to punish free-riders by expulsion if both of them are endowed with the expulsive ability. This is because the expulsion towards free-riders in a group can result in the enhancement of the proportion of contributors in this group. According to the aforementioned rules of the game, the payoff of the player on site $i$ obtained from the public goods game with bilateral costly expulsion that centred on, e.g. the player on site $j$, thus is

$$\left.\begin{aligned} P_{ij} &= \left(\frac{r n_{\mathrm{EC}}}{n_j} - 1\right) c - n_{\mathrm{ED}} c_E, \quad \text{if } S_i = \mathrm{EC} \\ \text{and} \qquad P_{ij} &= \frac{r n_{\mathrm{EC}} c}{n_j} - (n_{\mathrm{ED}} - 1) c_E, \quad \text{if } S_i = \mathrm{ED}, \end{aligned}\right\} \tag{2.1}$$

where the contribution amount $c$ in our model is set to 1 without loss of generality, $c_E \geq 0$ is the expulsion costs that are paid by expulsive players to banish each defector in the group, $n_j \in [2, k+1]$ denotes the size of the group that centred on the player on site $j$, and $n_{\mathrm{EC}}$ and $n_{\mathrm{ED}}$, satisfying $n_{\mathrm{EC}} + n_{\mathrm{ED}} = n_j$, represent the number of expulsive cooperators and that of expulsive defectors, respectively. Finally, every player has a chance to synchronously update its strategy by imitation. If the focal player at site $i$ randomly chooses an occupied site $j$, this player imitates the one at site $j$ with a probability:

$$F(P_j - P_i) = \frac{1}{1 + \exp[-\beta(P_j - P_i)]}, \tag{2.2}$$

where $\beta = 10$ quantifies the intensity of natural selection, which means that players with larger payoffs are readily imitated, although it is also possible to adopt the strategy of a player performing worse. Otherwise, the focal player at site $i$ will not change its strategy. In addition, we also introduce an extremely small rate of mutation $\mu = 0.00001$ into the strategy updating phase for the purpose of ensuring avoidance of the evolutionary system being stuck in artefact frozen states. Note that the negligible mutation rate does not change our main conclusion drawn in this article.

## 3. Results

To better appreciate the impacts of network structures in the public goods game in the presence of bilateral costly expulsion, it is useful to study evolutionary dynamics of the public goods game with bilateral costly expulsion in the mean field limit, which can be described by a replicator-like equation for the normalized density of expulsive cooperators $\bar{\rho}_{\mathrm{EC}} = \rho_{\mathrm{EC}}/\rho \in [0,1]$:

$$\frac{\partial \bar{\rho}_{\mathrm{EC}}}{\partial t} = (1-\mu)\bar{\rho}_{\mathrm{EC}}(1-\bar{\rho}_{\mathrm{EC}}) \tanh\left[\frac{\beta}{2}(\bar{P}_{\mathrm{EC}} - \bar{P}_{\mathrm{ED}})\right] + \mu(1-2\bar{\rho}_{\mathrm{EC}}), \tag{3.1}$$

where $\bar{P}_X (X \in \mathrm{EC}, \mathrm{ED})$ represents the expected payoffs for players (i.e. $\bar{P}_{\mathrm{EC}}$ for expulsive cooperators and $\bar{P}_{\mathrm{ED}}$ for expulsive defectors):

$$\left.\begin{aligned} \bar{P}_{\mathrm{EC}} &= \frac{(n-1)(r+nc_E)\bar{\rho}_{\mathrm{EC}}}{n} - (n-1)c_E + \frac{r}{n} - 1, \\ \text{and} \quad \bar{P}_{\mathrm{ED}} &= \frac{(n-1)(r+nc_E)\bar{\rho}_{\mathrm{EC}}}{n} - (n-1)c_E, \end{aligned}\right\} \tag{3.2}$$

where the group size $n$ should be set to $\rho(k+1)$ for convenience of comparison with the simulation results of the spatial public goods game with bilateral costly expulsion. From equation (3.2), we obtain $\bar{P}_{\mathrm{EC}} - \bar{P}_{\mathrm{ED}} = r/n - 1 < 0$. Therefore, the expected payoff difference between expulsive cooperators and expulsive defectors is independent on the variation of the expulsion cost $c_E$, and $\bar{\rho}_{\mathrm{EC}}$ tends to zero in the limit $\mu \to 0$ for arbitrary values of $c_E$ and $\beta$. In short, expulsion costs have neither beneficial nor detrimental effects on the evolution of public cooperation, and expulsive defectors become completely dominant in the mean field system.

In what follows, we first study the effects of bilateral costless expulsion in the spatial public goods game. To get a comprehensive insight, we compare the evolutionary outcomes for the whole applicable range of the enhancement factor $r$ as obtained with and without bilateral costless expulsion in figure 1. The presented results show that the bilateral costless expulsion does have beneficial effects on the evolution of public cooperation across the whole range of the enhancement factor $r$: bilateral costless expulsion promotes not only the emergence of but also the dominance of public cooperation (compare figure 1*a* with *c*). Moreover, we find that the coexistent states between cooperation and defection change from stationary to statistical ones when costless expulsion is introduced into the spatial public goods game in a bilateral manner (see figure 1*a*,*b*). Here, 'statistically coexistent state' means that the spatial population evolves to either defection dominated states or cooperation dominated states under the same condition of parameter settings. Put differently, expulsive cooperators and expulsive defectors coexist in the statistical sense (see figure 1*d*). Figure 2 reveals how the evolutionary fates of expulsive cooperators vary with the enhancement factor $r$ and the expulsion cost $c_E$. Surprisingly, we find that while expulsive cooperators and expulsive defectors may still statistically coexist but are impossible to stably coexist in the spatial population, expulsion costs can further support the evolution of cooperation in the spatial public goods game with bilateral expulsion. It should be noticed that, although the evolutionary mechanism of bilateral costless expulsion works here is, in principle, similar to that does in the spatial games of social dilemma governed by pairwise interactions, it still requires to clarify why and how the expulsion costs can further promote the evolution of cooperation in the spatial public goods game with bilateral expulsion.

For the purpose of revealing the constructive role of expulsion costs in the spatial dynamics of public cooperation, we conduct a mean field analysis of the public goods game with bilateral costly expulsion under the condition of heterogeneous interaction rates. Suppose that the probability of interaction among players is also dependent on their strategies [25]: a player interacts with other ones of the same kind with rate $w \in [0,1]$, and with other players adopting

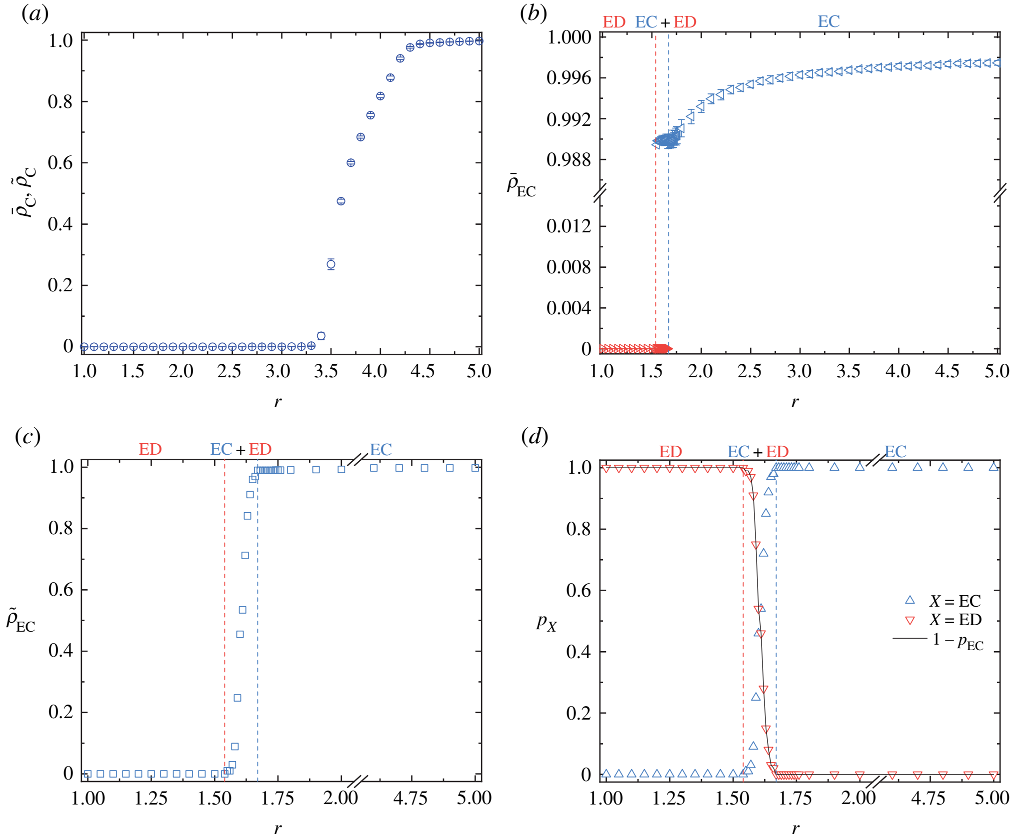


**Figure 1.** Effects of bilateral costless expulsion on the evolution of cooperation in the spatial public goods game. (*a*) Stationary fraction of cooperators $\bar{\rho}_C$ and average fraction of cooperators $\tilde{\rho}_C$ as a function of the enhancement factor $r$. (*b*) Stationary fraction of expulsive cooperators $\bar{\rho}_{\mathrm{EC}}$ in dependence on the enhancement factor $r$. Here, the spatial system evolves into either expulsive cooperators dominated states or expulsive defectors dominated states for the enhancement factor $r \in (1.54, 1.67)$, which belongs to the parameter area of statistically coexistent state. By comparing figure 1*a* with *b*, one can observe that the introduction of bilateral costless expulsion leads to a qualitative change of the spatial dynamics from stable coexistence between cooperators and defectors to statistical coexistence between expulsive cooperators and expulsive defectors in the respective parameter regions of coexistence. The error bars for stationary fractions of players in figure 1*a*,*b* mark one standard deviation. (*c*) Average fraction of expulsive cooperators $\tilde{\rho}_{\mathrm{EC}}$ in variation with the enhancement factor $r$. In comparison with figure 1*a*, we find that bilateral costless expulsion enhances the average level of cooperation for any value of the enhancement factor $r$. (*d*) Dominated probabilities of expulsive cooperators $p_{\mathrm{EC}}$ and expulsive defectors $p_{\mathrm{ED}}$ as a function of the enhancement factor $r$. The black solid line denotes the equation $p = 1 - p_{\mathrm{EC}}$, which coincides perfectly with the simulation results of $p_{\mathrm{ED}}$. Therefore, the spatial population is dominated by either expulsive cooperators or expulsive defectors in the public goods game with bilateral costless expulsion whatever the value of the enhancement factor is. This indicates that expulsive cooperators and expulsive defectors must statistically coexist but are not possible to stably coexist in the parameter regions associated with the coexistent states, as also illustrated by figure 1*b*. The two dashed vertical lines in figure 1*b*–*d* separate the final population states into three different classes: expulsive defectors dominated states (ED) satisfying $p_{\mathrm{EC}} = 0$, statistically coexistent states between expulsive cooperators and expulsive defectors (EC + ED) satisfying $p_{\mathrm{EC}} \in (0, 1)$ and expulsive cooperators dominated states (EC) satisfying $p_{\mathrm{EC}} = 1$, which are denoted along the top axis. Parameter settings: expulsion cost $c_E = 0$, system size $L \times L = 200 \times 200$ and population density $\rho = 0.9$. The final results are averaged over 100 independent runs with different random realizations. (Online version in colour.)

a different strategy with rate $1 - w$. Here, we make an implicit assumption that the tendency of an expulsive cooperator to interact with other expulsive cooperators is the same as that of an expulsive defector to interact with other expulsive defectors. Therefore, the expected payoffs for

royalsocietypublishing.org/journal/rspa Proc. R. Soc. A 477: 20210627


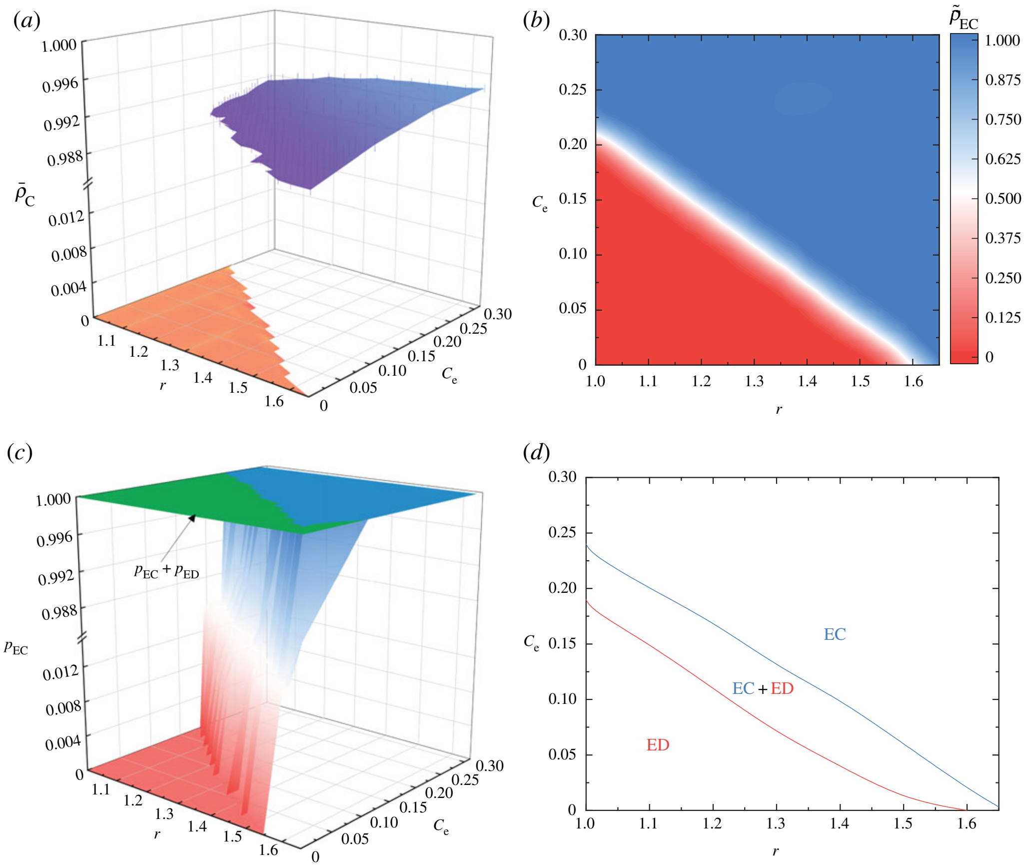


**Figure 2.** Impacts of expulsion costs in the evolution of public cooperation in the spatial population. (*a*) Stationary fraction of expulsive cooperators $\bar{\rho}_{\mathrm{EC}}$ as a function of the enhancement factor $r$ and the expulsion cost $c_E$. The error bars for stationary fractions of expulsive cooperators mark one standard deviation. Here, the spatial system evolves into either expulsive cooperators dominated states or expulsive defectors dominated states in the statistically coexistent state, the parameter region of which is shown more clearly in figure 1*d*. This means that the fundamental alteration of the evolutionary dynamics from stable to statistical coexistence by bilateral expulsion is robust against the variation of the expulsion cost $c_E$. (*b*) Average fraction of expulsive cooperators $\tilde{\rho}_{\mathrm{EC}}$ in dependence on the enhancement factor $r$ and the expulsion cost $c_E$. Here, we see clearly that the average level of cooperation is increased with the expulsion cost $c_E$. (*c*) Dominated probabilities of expulsive cooperators $p_{\mathrm{EC}}$ in variation with the enhancement factor $r$ and the expulsion cost $c_E$. The upmost plane represents the sum of the dominated probability of expulsive cooperators and that of expulsive defectors $p_{\mathrm{EC}} + p_{\mathrm{ED}}$ in variation with the enhancement factor $r$ and the expulsion cost $c_E$. Note that $p_{\mathrm{EC}} + p_{\mathrm{ED}}$ in this plane is always equal to one, which indicates that the spatial population is dominated by either expulsive cooperators or expulsive defectors in the parameter areas associated with the coexistent states, and thus that expulsive cooperators and expulsive defectors may statistically coexist but are impossible to stably coexist in the spatial public goods game. (*d*) Phase diagram on the $r - c_E$ parameter plane, depicting three different parameter areas satisfying $p_{\mathrm{EC}} = 0$ (i.e. the $\mathrm{ED}$ phase), $p_{\mathrm{EC}} \in (0, 1)$ (i.e. the $\mathrm{EC} + \mathrm{ED}$ phase) and $p_{\mathrm{EC}} = 1$ (i.e. the $\mathrm{EC}$ phase), respectively. Parameter settings: system size $L \times L = 200 \times 200$ and population density $\rho = 0.9$. The final results are averaged over 100 independent runs with different random realizations. (Online version in colour.)

players with non-uniform interaction rates become

$$\left.\begin{aligned}\bar{P}_{\mathrm{EC}} &= \frac{(n-1)(r+nc_E)}{n}\frac{w\bar{\rho}_{\mathrm{EC}}}{w\bar{\rho}_{\mathrm{EC}}+(1-w)(1-\bar{\rho}_{\mathrm{EC}})} - (n-1)c_E + \frac{r}{n} - 1,\\ \text{and}\quad \bar{P}_{\mathrm{ED}} &= \frac{(n-1)(r+nc_E)}{n}\frac{(1-w)\bar{\rho}_{\mathrm{EC}}}{(1-w)\bar{\rho}_{\mathrm{EC}}+w(1-\bar{\rho}_{\mathrm{EC}})} - (n-1)c_E,\end{aligned}\right\} \tag{3.3}$$

where the group size $n$ is set to $\rho(k+1)$. For equation (3.3), we offer the following observations: if $w > 1/2$, then players prefer to interact with partners who use the same strategy as them; if, however, $w < 1/2$, then players are more likely to interact with co-players who adopt a different strategy from them; otherwise (i.e. $w = 1/2$), players interact with other ones indiscriminately, and equation (3.3) is reduced to the expected payoffs for players under the condition of uniform interaction rates, i.e. equation (3.2). For the case of heterogeneous interaction rates, the difference between the expected payoff of expulsive cooperators and that of expulsive defectors can be given as follows:

$$\bar{P}_{\mathrm{EC}} - \bar{P}_{\mathrm{ED}} = \frac{(n-1)\bar{\rho}_{\mathrm{EC}}(1-\bar{\rho}_{\mathrm{EC}})(2w-1)(r+nc_E)}{n[w\bar{\rho}_{\mathrm{EC}} + (1-w)(1-\bar{\rho}_{\mathrm{EC}})][(1-w)\bar{\rho}_{\mathrm{EC}} + w(1-\bar{\rho}_{\mathrm{EC}})]} + \frac{r}{n} - 1. \tag{3.4}$$

For imitation dynamics of spatial games, $w > 1/2$ is satisfied because players of the same strategy are self-organized into spatial clusters [26]. Hence, we have $2w - 1 > 0$, which indicates that $\bar{P}_{\mathrm{EC}} - \bar{P}_{\mathrm{ED}}$ is increased with the expulsion cost $c_E$ (see equation (3.4)). On the other hand, let the polynomial $G(\bar{\rho}_{\mathrm{EC}})$ be

$$\begin{aligned} G(\bar{\rho}_{\mathrm{EC}}) = (2w-1)[(n-r)(2w-1) - (n-1)(r+nc_E)]\bar{\rho}_{\mathrm{EC}}^2 \\ - (2w-1)[(n-r)(2w-1) - (n-1)(r+nc_E)]\bar{\rho}_{\mathrm{EC}} - (n-r)w(1-w). \end{aligned} \tag{3.5}$$

Note that $G(\bar{\rho}_{\mathrm{EC}})$ has the same shapes of function with the payoff difference in equation (3.4) and thus determines the non-trivial equilibria of the deterministic dynamics described by equation (3.1) in the limit $\mu \to 0$. As $w \in (1/2, 1]$, $c_E \geq 0$ and $r \in (1, n)$, we have $(2w-1)[(n-r)(2w-1) - (n-1)(r+nc_E)] < 0$ meaning that $G(\bar{\rho}_{\mathrm{EC}})$ attains its maximum value $G(\bar{\rho}_{\mathrm{EC}})_{\max}$ at $\bar{\rho}_{\mathrm{EC}} = 1/2$:

$$G(\bar{\rho}_{\mathrm{EC}})_{\max} = \frac{(n-1)(2w-1)(r+nc_E) - (n-r)}{4}. \tag{3.6}$$

Depending on the sign of $G(\bar{\rho}_{\mathrm{EC}})_{\max}$, we can classify the mean field dynamics into the following three categories:

(i) if $c_E < \bar{c}_E = (n-r)/(n(n-1)(2w-1)) - (r/n)$, ED completely dominates EC;
(ii) if $c_E = \bar{c}_E = (n-r)/(n(n-1)(2w-1)) - (r/n)$, ED dominates EC, and there exists one unstable interior fixed point locating at $\bar{\rho}_{\mathrm{EC}} = 1/2$;
(iii) if $c_E > \bar{c}_E = (n-r)/(n(n-1)(2w-1)) - (r/n)$, there exist two interior fixed points with the left one being unstable and the right one being stable as well as two fixed points on the boundary with ED being stable and EC being unstable.

Therefore, the state of expulsive cooperators can transition from extinction to survival or even to domination as expulsion cost $c_E$ increases from its value below $\bar{c}_E$ to that beyond $\bar{c}_E$. For both case (i) and case (ii), the increase of expulsion cost $c_E$ can reduce the absolute value of the expected payoff difference between expulsive cooperators and expulsive defectors $|\bar{P}_{\mathrm{EC}} - \bar{P}_{\mathrm{ED}}|$ though expulsive cooperation is inevitably dominated by expulsive defection (i.e. $\bar{P}_{\mathrm{EC}} - \bar{P}_{ED} \leq 0$ for both case (i) and case (ii)]. For case (iii), we obtain $G(\bar{\rho}_{\mathrm{EC}})_{\max} > 0$ because $c_E > (n-r)/(n(n-1)(2w-1)) - (r/n)$. Considering this result together with the condition that $G(0) = -(n-r)w(1-w) < 0$ and $G(1) = -(n-r)w(1-w) < 0$, one can conclude that there exist two roots for equation (3.5) with the left one $\bar{\rho}^L_{\mathrm{EC}} \in [0, 1/2)$ as well as the right one $\bar{\rho}^R_{\mathrm{EC}} \in (1/2, 1]$. As $\partial G(\bar{\rho}_{\mathrm{EC}})/\partial \bar{\rho}_{\mathrm{EC}} = [(n-r)(2w-1)^2 - (n-1)(2w-1)(r+nc_E)](2\bar{\rho}_{\mathrm{EC}} - 1)$, $G(\bar{\rho}_{\mathrm{EC}})$ increases monotonically with $\bar{\rho}_{\mathrm{EC}} \in [0, 1/2]$ and decreases monotonically with $\bar{\rho}_{\mathrm{EC}} \in [1/2, 1]$. Therefore, there exist two interior fixed points with the left one $\bar{\rho}^L_{\mathrm{EC}}$ being unstable and the right one $\bar{\rho}^R_{\mathrm{EC}}$ being stable for the deterministic dynamics described by equations (3.1) and (3.3). Because the left fixed point $\bar{\rho}^L_{\mathrm{EC}}$ and the right one $\bar{\rho}^R_{\mathrm{EC}}$ are symmetric with respect to $\bar{\rho}_{\mathrm{EC}} = 1/2$, we can choose the Euclidean distance between $\bar{\rho}_{\mathrm{EC}}^{L}$ and $\bar{\rho}_{\mathrm{EC}}^{R}$,

$$d = |\bar{\rho}_{\mathrm{EC}}^{R} - \bar{\rho}_{\mathrm{EC}}^{L}| = \sqrt{1 - \frac{4(n-r)w(1-w)}{(2w-1)[(n-1)(r+nc_E)-(n-r)(2w-1)]}}, \tag{3.7}$$

as a measurement to determine how expulsion costs affect the evolutionary dynamics in case (iii). From equation (3.7), one can find that $d$ is increased with expulsion cost $c_E$, which indicates that the increment of $c_E$ leads to both the reduction of invasion barrier $\bar{\rho}_{\mathrm{EC}}^{L}$ for expulsive cooperators and the enhancement of the final cooperation level at the stable fixed point $\bar{\rho}_{\mathrm{EC}}^{R}$ once the frequency of expulsive cooperators exceeds the invasion barrier. Interestingly, the latter argument can be confirmed by the simulation results of the spatial public goods game with bilateral costly expulsion presented in figure 2*a*.

## 4. Discussion

In conclusion, we have shown that expulsion is still beneficial for the evolution of public cooperation though it is assigned to both cooperators and defectors of the spatial public goods game in our model. Furthermore, the bilateral introduction of expulsion to the spatial public goods game results in the interesting alteration of evolutionary dynamics: the emergence of statistically coexistent states where the spatial population is dominated by either expulsive cooperators or expulsive defectors, and the disappear of coexistent steady states where cooperators and defectors stably coexist at the interior fixed points of the spatial system. Importantly, expulsion costs, which are irrelevant for the standard mean field dynamics under the condition of uniform interaction rates, play an unexpectedly constructive role in the cooperative dynamics of the spatial public goods game with bilateral expulsion. The positive assortment among players with the same strategy arising in the imitation dynamics of spatial games leads to the result that expulsive defectors have a higher frequency to interact with their counterparts than expulsive cooperators do [27]. Therefore, expulsive defectors, on average, bear more expulsion costs than their evolutionary competitors do, which in turn results in the further promotion of public cooperation by bilateral expulsion. The validity of this argument is confirmed by a mean field analysis of the public goods game with bilateral costly expulsion under the condition of heterogeneous interaction rates. In fact, the impacts of other costly other-regarding behaviours in the evolution of public cooperation have also been investigated in recent years. Of these behaviours, the most relevant class to costly expulsion is costly punishment [28]. By costly punishment, it usually means that punishers can pay a cost to impose a larger cost on defective interaction partners [29,30]. Generally speaking, costly expulsion in our work can be regarded as a special kind of costly punishment. Particularly, Helbing *et al.* have studied the bilateral effects of costly punishment on the evolution of cooperation in the spatial public goods game, and have found that punishment costs are negatively related to public cooperation though costly punishment can play a positive role in the evolution of cooperation in the spatial populations [31]. In comparison, our study shows that expulsion costs can even be beneficial for the evolution of public cooperation through spatial interactions.

Finally, it should be noted that bilateral costly expulsion may also be helpful to resolve the dilemma of voluntary vaccination [32]. In this case, individuals of a population can be classified into two categories from the viewpoint of evolutionary game theory: vaccinated individuals and unvaccinated individuals (i.e. free-riders). Here, free-riders refer to the individuals who refuse to get vaccinated but try to keep healthy by successfully exploiting the vaccination efforts of others. In such a voluntary vaccination dilemma game, the expulsive behaviours can be performed by both the vaccinated individuals and the unvaccinated individuals, personally or collectively, towards the free-riders for the purpose of enjoying the public goods of vaccination in the largest degree. In the well-mixed population, bilateral costly expulsion has no effect on the vaccination coverage because it cannot alter the uniform interaction rates between vaccinated individuals and free-riders. In the graph-structured population, however, one can expect the level of vaccination coverage can be enhanced by bilateral costly expulsion due to the heterogeneous interaction rates resulted from imitation dynamics of the graph-structured voluntary vaccination dilemma game though it requires further confirmation by theoretical evidence as well as simulation results.

Data accessibility. This article has no additional data.

Authors' contributions. X.W. and M.P.: designed and performed the research as well as wrote the paper.

Competing interests. We declare we have no competing interests.

Funding. X.W. acknowledges funding from the National Natural Science Foundation of China (grant no. 61903077), the Shanghai Sailing Program (grant no. 19YF1402500) and the Fundamental Research Funds for the Central Universities (grant no. 2232019D3-56). M.P. acknowledges funding from the Slovenian Research Agency (grant nos. P1-0403, J1-2457 and J1-9112).

Acknowledgements. X.W. acknowledges funding from the National Natural Science Foundation of China (Grant no. 61903077), the Shanghai Sailing Program (Grant no. 19YF1402500), and the Fundamental Research Funds for the Central Universities (Grant no. 2232019D3-56). M.P. acknowledges funding from the Slovenian Research Agency (Grant nos P1-0403, J1-2457, and J1-9112).

## References

1. Nowak MA. 2006 *Evolutionary dynamics: exploring the equation of life*. Cambridge, MA: Harvard University Press.
2. Lieberman E, Hauert C, Nowak MA. 2005 Evolutionary dynamics on graphs. *Nature* **433**, 312–316. (doi:10.1038/nature03204)
3. Santos FC, Pacheco JM. 2005 Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104. (doi:10.1103/PhysRevLett.95.098104)
4. Ohtsuki H, Hauert C, Lieberman E, Nowak MA. 2006 A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505. (doi:10.1038/nature04605)
5. Masuda N. 2007 Participation costs dismiss the advantage of heterogeneous networks in evolution of cooperation. *Proc. R. Soc. B* **274**, 1815–1821. (doi:10.1098/rspb.2007.0294)
6. Wang Z, Kokubo S, Jusup M, Tanimoto J. 2015 Universal scaling for the dilemma strength in evolutionary games. *Phys. Life Rev.* **14**, 1–30. (doi:10.1016/j.plrev.2015.04.033)
7. Allen B, Lippner G, Chen Y, Fotouhi B, Momeni N, Yau ST, Nowak MA. 2017 Evolutionary dynamics on any population structure. *Nature* **544**, 227–230. (doi:10.1038/nature21723)
8. Ito H, Tanimoto J. 2018 Scaling the phase-planes of social dilemma strengths shows game-class changes in the five rules governing the evolution of cooperation. *R. Soc. Open Sci.* **5**, 181085. (doi:10.1098/rsos.181085)
9. Fotouhi B, Momeni N, Allen B, Nowak MA. 2019 Evolution of cooperation on large networks with community structure. *J. R. Soc. Interface* **16**, 20180677. (doi:10.1098/rsif.2018.0677)
10. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. 2006 Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308. (doi:10.1016/j.physrep.2005.10.009)
11. Estrada E. 2012 *The structure of complex networks: theory and applications*. Oxford, UK: Oxford University Press.
12. Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A, Young JG, Petri G. 2020 Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* **874**, 1–92. (doi:10.1016/j.physrep.2020.05.004)
13. Perc M, Gómez-Gardeñes J, Szolnoki A, Floría LM, Moreno Y. 2013 Evolutionary dynamics of group interactions on structured populations: a review. *J. R. Soc. Interface* **10**, 20120997. (doi:10.1098/rsif.2012.0997)
14. Dawes RM. 1980 Social dilemmas. *Ann. Rev. Psychol.* **31**, 169–193. (doi:10.1146/annurev.ps.31.020180.001125)
15. Sigmund K. 2010 *The calculus of selfishness*. Princeton, NJ: Princeton University Press.
16. Ginsberg AG, Fu F. 2019 Evolution of cooperation in public goods games with stochastic opting-out. *Games* **10**, 1–27. (doi:10.3390/g10010001)
17. Santos FC, Santos MD, Pacheco JM. 2008 Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454**, 213–216. (doi:10.1038/nature06940)
18. Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. 2017 Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51. (doi:10.1016/j.physrep.2017.05.004)
19. Szolnoki A, Perc M, Szabó G. 2009 Topology-independent impact of noise on cooperation in spatial public goods games. *Phys. Rev. E* **80**, 056109. (doi:10.1103/PhysRevE.80.056109)

20. Perc M, Szolnoki A. 2010 Coevolutionary games—a mini review. *BioSystems* **99**, 109–125. (doi:10.1016/j.biosystems.2009.10.003)
21. Helbing D, Yu W. 2009 The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc. Natl Acad. Sci. USA* **106**, 3680–3685. (doi:10.1073/pnas.0811503106)
22. Wu T, Fu F, Wang L. 2011 Moving away from nasty encounters enhances cooperation in ecological prisoner's dilemma game. *PLoS ONE* **6**, e27669. (doi:10.1371/journal.pone.0027669)
23. Wu T, Fu F, Wang L. 2018 Coevolutionary dynamics of aspiration and strategy in spatial repeated public goods games. *New J. Phys.* **20**, 063007. (doi:10.1088/1367-2630/aac687)
24. Wang X, Duh M, Perc M. 2020 Robust cooperation against mutations via costly expulsion. *EPL* **132**, 38001. (doi:10.1209/0295-5075/132/38001)
25. Taylor C, Nowak MA. 2006 Evolutionary game dynamics with non-uniform interaction rates. *Theor. Popul. Biol.* **69**, 243–252. (doi:10.1016/j.tpb.2005.06.009)
26. Nowak MA, May RM. 1992 Evolutionary games and spatial chaos. *Nature* **359**, 826–829. (doi:10.1038/359826a0)
27. Langer P, Nowak MA, Hauert C. 2008 Spatial invasion of cooperation. *J. Theor. Biol.* **250**, 634–641. (doi:10.1016/j.jtbi.2007.11.002)
28. Sigmund K. 2007 Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22**, 593–600. (doi:10.1016/j.tree.2007.06.012)
29. Fehr E, Gächter S. 2002 Altruistic punishment in humans. *Nature* **415**, 137–140. (doi:10.1038/415137a)
30. Espín AM, Brañas-Garza P, Herrmann B, Gamella JF. 2012 Patient and impatient punishers of free-riders. *Proc. R. Soc. B* **279**, 4923–4928. (doi:10.1098/rspb.2012.2043)
31. Helbing D, Szolnoki A, Perc M, Szabó G. 2010 Evolutionary establishment of moral and double moral standards through spatial interactions. *PLoS Comput. Biol.* **6**, e1000758. (doi:10.1371/journal.pcbi.1000758)
32. Chen X, Fu F. 2019 Imperfect vaccine and hysteresis. *Proc. R. Soc. B* **286**, 20182406. (doi:10.1098/rspb.2018.2406)