# Nonconformity of cooperators promotes the emergence of pure altruism in tag-based multi-agent networked systems

Tarik Hadzibeganovic [a], Pengbi Cui [b,c,*], Zhi-Xi Wu [d,e]

[a] Department of Psychology, Faculty of Natural Sciences, University of Graz, 8010 Graz, Austria
[b] Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China
[c] Big Data Research Center, University of Electronic Science and Technology of China, Chengdu 611731, China
[d] Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou Gansu 730000, China
[e] Key Laboratory for Magnetism and Magnetic Materials of the Ministry of Education, Lanzhou University, Lanzhou Gansu 730000, China

## ABSTRACT

Designing cooperation-enhancing protocols for large-scale multiagent networked systems has been a grand transdisciplinary challenge. In recent years, tag-based interactions and conformity bias have been studied extensively but separately as two viable mechanisms for cooperation enhancement in such systems. Inspired by recent studies on interaction effects in social dilemmas, we herein develop a hybrid, multiagent-based, co-evolutionary model of tag-mediated cooperation and conformity with conditional and unconditional strategies. Through a series of extensive Monte Carlo simulation experiments, we study four variants of this computational model, finding that under the majority rule, the nonconforming unconditional cooperators and conformity biased transmission of other strategies can lead to global altruistic dominance. Employing a random pinning control mechanism, we further observe that only a small fraction of nonconforming altruists is actually required to drive the system towards a robust persistence of pure altruism. Our analytic results in combination with further computational experiments reveal that spatial structure and nonconformity of cooperators are the two indispensable ingredients for the stable dominance of altruistic behavior in tag-based multiagent systems. Our findings can be beneficial for developing novel cooperation-controlling techniques in distributed self-organizing systems such as peer-to-peer networks or in various social networking and viral marketing technologies.

## 1. Introduction

Establishing cooperation and coordination in large-scale, decentralized, and dynamic multiagent systems such as peer-to-peer (P2P) networks has been a grand transdisciplinary challenge [1–9]. In recent years, tag-mediated interactions have been explored as a potential mechanism for attaining sustainable cooperative outcomes in such systems [10,11]. In models of tag-based cooperative behavior [12–24], interacting components harness the shared mutual similarity level to inform their decision making, such that benevolent actions are directed either towards sufficiently similar [24] or towards identical others [12].

Tag-based interactions are cognitively effortless [24], computationally undemanding [10], and strikingly robust against parameter variations [12] as they rely on shallow memoryless processing of partial information that is accessible from the surrounding environment. Consequently, the memory of previous interactions or the relatedness among interactors is not a necessary prerequisite for the outbreak of cooperation in tag-based systems, thereby making them particularly useful in decentralized large-scale networks in which reputation [25] or trust [4] as cooperation-promoting mechanisms are unreliable due to multiple anonymous, unrepeated interactions [24].

However, in spite of rather comprehensive previous research, earlier tag-based cooperation models did not explicitly consider the underlying mechanisms that actually cause individuals to cooperate with similar tag-mates [14]. For example, it has widely been assumed that the preference for ingroup interactions and cooperative exchanges within the group were largely associated with the perceived shared similarity among interactors [26]. This localist view of previous computational models is, in part, understandable, given the fact that within the framework of agent-based modeling and simulation, they addressed the global emergence of tag-mediated cooperation out of local interactions among agents who did not have the understanding of the group concept or the explicit knowledge of any of its underlying properties. On the other hand, it is reasonable to assume that individuals can cooperate not just due to the local processes of tag-recognition and the associated

* Corresponding author at: Web Sciences Center, University of Electronic Science and Technology of China, Chengdu 611731, China.
*E-mail addresses:* tarik.hadzibeganovic@gmail.com (T. Hadzibeganovic), cuisir610@gmail.com (P. Cui), wuzhx@lzu.edu.cn (Z.-X. Wu).

tag-mediated affinity towards similar others, but also due to global processes such as the ongoing pressures to conform to collectively desirable group norms [27–29].

Notably, conformity [30–35] itself has attracted much recent attention in cooperation research [36–43], yielding rather mixed findings about its role as a catalyst of cooperative behavior [42–46]. For example, cooperation is more likely to occur if an individual is facing a cooperative partner and an uncooperative group than when meeting an uncooperative partner and a cooperative group, suggesting thus that reciprocity is superior to conformity in promoting cooperative venture [42]. Moreover, it has been shown that the coevolutionary dynamics of cooperation and conformity alone may not explain the observed levels of cooperation in human societies, though conformist bias may still have substantial cooperation-enforcing potency when combined with other factors [43].

Conformity may thus exhibit nontrivial interaction effects such that different outcomes will be observed when conformity is combined with varying levels of the same factor. For example, conformity knowingly interacts with the type of game that is being played among interacting agents. Naturally, in coordination games [47], conforming to the opinion of the majority is beneficial for the individual, whereas in complementarity games [48], countering the opinion of the majority is the better payoff-maximizing option (also see Refs. [44,49]).

Intriguingly, even though conformity appears to better support cooperation in the presence of other mechanisms and could be associated with the ubiquity of ingroup biased behaviors, it has never been investigated previously in combination with tag-based cooperation [12], in which prosocial behavior is typically mediated via tags, phenotypic features that serve as discernible markers of group membership and signals of cooperative potential among group members. Instead, studies of tag-based interactions and conformity bias have traditionally been carried out separately, showing that none of the two mechanisms alone can push the system into the state of stable pure altruism that is desirable in many large-scale, decentralized, complex networked systems.

However, as an integral part of signaling mechanisms [22], tags and tag-based interactions are ubiquitous in social systems in which they often shape the altruistic norm compliance and norm imposition processes that are usually constrained within groups [50]. Tag-based cooperation models have initially been developed to understand socio-biological and behavioral phenomena such as ethnocentrism [12], in-group favoritism [51], or greenbeard chromodynamics [13,15]. Related models with tag-mediated behavior have been employed to explain indirect reciprocity [14, 24] as a viable mechanism for the evolution of cooperation. Among the most common findings of these earlier modeling studies is the superiority of the conditional 'ethnocentric' cooperation – generosity directed exclusively towards tag-sharing group members – that evolves and robustly prevails across a wide range of conditions. However, since evolutionary games with conditional strategies can theoretically exhibit multiple alternative equilibria (see e.g. [22]), recent studies [18,19] have initiated the quest for alternative outcomes in tag-based cooperation models, especially those with potential technological applications [11].

For example, tag-based interactions have been used for cooperation support in file-sharing P2P systems [10], where it is essential to prevent self-interested behavior in which free-riding peers download files from other generous users without contributing with their own file uploads. However, cooperation in tag-based systems is additionally vulnerable to exploitation and subject to collapse if individual nodes are capable of faking the tag to receive the benefits, while simultaneously refusing to provide any costly donations in return. Due to this challenging difficulty, it is necessary to implement efficient cheater detection and neutralization

mechanisms, e.g. to dynamically change the cooperative tag phenotypes before they can be adopted by the free-riders.

Motivated by these limitations of earlier approaches, and inspired by recent studies on interaction effects in social dilemmas [23,52–54], we attempted to elucidate the combined effects of conformity and tags with other potential mechanisms such as spatial structure that could give rise to sustainable cooperative outcomes in multiagent networked systems. Critically, since tag-mediated interactions typically lead to the dominance of the intra-group conditional cooperation (that outcompetes all other strategies), we addressed the question of whether and under what circumstances global pure altruism can be reached in the presence of conformity in a tag-based multi-agent system, i.e. a system characterized by tag-mediated interactions among multiple computational agents who condition their cooperative behavior on the observable phenotypic 'tags' of their opponents. To this effect, we employed the methods of game theory, evolutionary computing, and agent-based simulation [55–61] to institute a hybrid tag-mediated cooperation model with an implemented conformity bias.

The multi-agent framework has attracted much recent attention across disciplines, and has been identified as one of the central research hotspots in the knowledge-based systems literature [62]. In particular, it has been beneficial for studying social behaviors in networked systems with different interaction topologies, where spatial structure is crucial for understanding the investigated phenomenon, when agents' positions in this structure are not fixed, and when agents and their interactions are heterogeneous. Multiagent systems can thus flexibly simulate complex architectures of real-world social networks, their long-term evolutionary dynamics, and the underlying diversity of contact strengths which can coevolve with topological network properties. Moreover, they can serve as powerful forecasting and decision support tools in complex networked systems, e.g. for traffic, epidemic, or operational risk management [55].

Multiagent-based models are known for their robustness and tractability of low-cost solutions under more realistic scenarios without significant loss of rigor [55,59]. Perhaps most importantly, they can adequately capture emergent nonlinear phenomena arising from temporal instabilities and spatial correlations [63] that are otherwise neglected in equation-based macroscopic models. Specifically, in our study, using multiagent networked systems allows us to identify and systematically isolate the effects of spatial structure from those of conformity on the evolution of strategies in tag-based systems, and to better understand the underlying spatial pattern-formation processes.

To examine how conformity affects cooperation in tag-based multiagent networked systems, we systematically investigated four different variants of our hybrid computational model. We performed a series of extensive Monte Carlo simulations and we conducted mathematical analyses and further numerical experiments to identify the specific altruism-promoting conditions. In addition, we introduced several new measures to enhance the interpretation of our results and explain the underlying differences between the competing strategies in our model.

Thus, different from most previous studies of tag-based cooperation, the focus of the present paper was set on understanding the evolutionary dynamics of altruism, i.e. of unconditional cooperation, and whether and how it can emerge and outcompete other conditional and unconditional strategies in the presence of conformity. In this respect, this is to our knowledge the first work on the role of conformity in tag-based cooperative systems [12]. Moreover, our study is the first to investigate a specific mechanism for the control of global altruism and its robustness in tag-based multi-agent networks, with potential applications to various technological systems. We also note that previous models addressing

the role of conformity in cooperation were largely limited to the implementation of simple unconditional strategies (e.g. "always cooperate" and "always defect"), but they have never addressed the potentials of combining the conformity mechanism with both tag-based conditional and standard unconditional strategies that were for the first time examined within the context of our present model.

Importantly, our model considers the inheritance of conformity traits and the conformity biased transmission of strategies in the context of an explicit co-evolutionary model with tag-mediated interactions; by contrast, to the best of our knowledge, previous models of the role of conformity in cooperation did not address the *evolution* but only the *development* and learning of strategies under the influence of conformity within the same, fixed population of individuals. Instead, our model considers an *open* [5] and *reproducing* population of artificial decision makers subject to mutation and birth–death dynamics [12,64] with a fluctuating population size.

We have organized the remainder of our article as follows. In Section 2 (Methods), we describe the overall features of our model, the conformity mechanism, and the evolutionary algorithm underlying our simulation experiments. In Section 3, we present and discuss the main findings of our computational experiments. These findings are presented in four subsections, separately for each of the four model variants. The presentation of the results for each model variant is always preceded by a brief overview of its main features and comparisons with other model variants. A general discussion of our main findings, future research directions, and potential applications are then given in Section 4. We conclude with a summary of the main results and a brief discussion of their overall importance in Section 5. Our further exploration of the parameter space, with additional computational experiments that verified the robustness of our results, is presented in Appendix A. Finally, the results of our analytic approach and their systematic comparisons with numerical simulations are given in Appendix B.

## 2. Methods

### 2.1. General overview of the model

The main algorithm underlying our current evolutionary model was inspired by the classic Hammond–Axelrod (HA) study on tag-mediated ethnocentric cooperation [12], which is now extended by including the conformity mechanism [37]. This mechanism introduces a novel dimension into the model, namely, that of cooperative and self-interested preferences of an individual, which can be modified by the preferences of the group, and which in turn can determine an agent's ultimate behavior (cooperation vs. defection) given the phenotypes (tags) of its competitors. Moreover, unlike the classic model [12], which only studied the evolution of strategies in a simple regular square lattice, we study the co-evolutionary dynamics of strategies and conformity on Watts–Strogatz small-world networks [65] with varying degrees of network rewiring, which enables us to address the potential interaction effects of population structure and conformity on cooperation. Thus, heritable conformity traits co-evolve in our tag-based cooperation model with the conformity-biased conditional and unconditional strategies in a small-world networked system.

In the following two subsections, we first provide an overview and the main characteristics of our conformity mechanism, and we then explain how the conformity mechanism is actually implemented into the evolutionary process of our tag-based cooperation model.

### 2.2. Properties of the conformity mechanism

In our model, conformity itself does not change the strategies or tags of individuals. Instead, conformity affects an individual's preference or willingness to cooperate that can be changed by the states of the surrounding individuals. We distinguish between the cooperative preference ('Yes') and the self-interested preference ('No').

With some probability, each individual $i$ can turn on its cooperative preference (i.e. the preference either shifts to or remains in the cooperative state 'Yes'). The exact form of the conformity probability function in our model is given by [37]:

$$\Phi(\omega_{s_j}) = \frac{1}{1 + \exp(-\alpha(\omega_{s_j} - \delta))}, \qquad \omega_{s_j} = m_c/m \qquad (1)$$

where $m_c$ is the number of group members with the cooperative preference 'Yes' and $m$ represents the total number of individuals within the same group of the $l$-step nearest neighbors, excluding the focal individual $i$. Thus, the cooperative preference of a player $i$ does not contribute to the value of $m_c$. The parameter $\alpha$ denotes the *conversion sensitivity*, i.e. an agent's sensitivity to the states of the surrounding group members or simply the probability that an agent will convert its current preference to that of the group. The parameter $\delta$ stands for the *strength of influence* that is exerted by group members on an individual. It is selected as a threshold-value parameter $\delta \in [0\ 1.0)$ to weigh the strength of group influence on an individual agent. Lower values of $\delta$ typically lead to a stronger influence in the model, such that only a few neighbors sharing the same preference can cause the focal player to switch the preference to that of the influencing group. Conversely, larger values of $\delta$ correspond to weaker influence, where more neighbors with the same preference are necessary to successfully exert an influence on a targeted, focal individual. Thus, in our current model, conformity follows the *minority rule* when $\delta < 0.5$, whereas the *majority rule* is in action when $\delta > 0.5$.

If an individual turns on its cooperative preference as a result of conformity, it will cooperate with its designated nearest neighbors if and only if the tag-based model conditions are also satisfied. For example, in the standard tag-based cooperation model [12], there are four types of strategists: Ethnocentric, altruistic, egoistic, and cosmopolitan. The cosmopolitan agents cooperate only with the opponents having a different tag color; otherwise, they always defect. Thus, in our tag-based model with conformity, one cosmopolitan agent with cooperative preference 'Yes' will cooperate with its neighbors if and only if these neighbors are alive and their tag colors are different. In addition, the connection between them has to be a local one, i.e., agents cannot cooperate via long-range connections in the basic tag-based model with conformity (in our further model variants, we additionally study the influence of long-range connections, where interactions are enabled also via long-range network links). On the other hand, cosmopolitans with the self-interested preference 'No' will never cooperate (as long as they carry this preference), regardless of the types of connections (local vs. global), the living states (alive or dead), or the tag colors of their opponents.

Since we employ birth–death dynamics in our computational model, at each time step, an agent can be randomly selected to die (for details, please see the next subsection with our evolutionary algorithm). This further means that following its death, an individual agent disappears from the network and creates an empty space for the new offspring or other agents that can invade the system from the outside. Therefore, in our computational model, agent interactions naturally occur among 'living' agents that actually exist in the system and occupy the network nodes, but not between the empty nodes which are not populated. On the other hand, in our analytic model (see Appendix B), we used an aspatial model
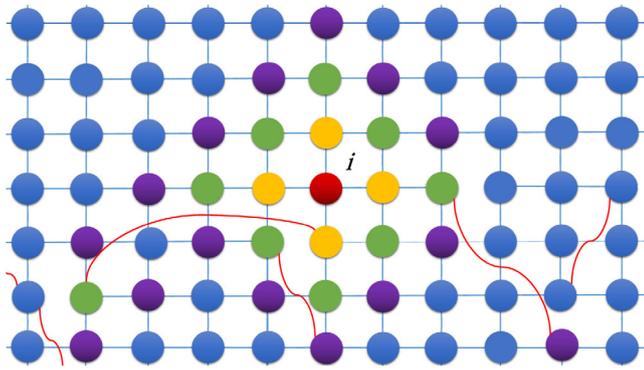
**Fig. 1.** Types of interaction neighborhood in a 2d Watts–Strogatz (WS) small-world network. The local connections are depicted as blue links, whereas the long-range connections are represented via red links. The first nearest neighbors ($l = 1$) are depicted as orange nodes surrounding the focal, red node $i$. For $l = 2$, the group focusing on $i$ contains both the orange and the green nodes, respectively. For example, the leftmost green node in the bottom-left corner is the long-range second-nearest neighbor of the node $i$. The interaction neighborhood $l = 3$ consists of orange, green, and violet nodes focusing on the central, red node $i$.

version in which the birth–death process is not applied and the population size is therefore constant. In such cases, all agents are referred to as 'living' and are all equally likely to interact with everyone else in the population. Thus, the restriction of spatial model variants (in which only 'living' agents can interact with each other) is not present in such analytic model variants.

As we consider several different variants of our model, it is necessary to note that in our model simulations and in the analysis (see Appendix B), we distinguish between the *nonconformist (free) altruists* and *conformist altruists*. The preferences of the nonconformist free altruists are not influenced by the preferences of their surrounding group members and are therefore always set as 'Yes', whereas the preferences of conformist altruists are susceptible to conformity bias.

Since they are unconditional defectors, it is natural to assume that the actual behavior of egoists cannot be affected by the conformity bias in the present model. In other words, egoists can change their own preferences due to group influence but their actual behavior (unconditional defection) remains unchanged throughout the simulation; in addition, egoists can exert the influence on others via their own preferences. Initially, the preferences of all individuals except egoists are other-regarding (i.e. they have the cooperative preference state 'Yes').

Every agent interacts only with its connected nearest neighbors, as specified by the interaction neighborhood size in the model (see Fig. 1). More specifically, every agent can only give donations to or receive benefits from its connected nearest neighbors (i.e., its 1st nearest neighbors), whereas its cooperative preference is under the influence of those group members that are located within the $l$th nearest neighbors, in accordance with the conformity rule given by the Eq. (1). As depicted in Fig. 1, the second-nearest neighbors are the nearest connected neighbors of the first-nearest neighbors of the central node $i$ (excluding both the central node $i$ and its first-nearest neighbors), and similarly, the third-nearest neighbors are the nearest connected neighbors of the second-nearest neighbors of the central node $i$ (excluding again the central node $i$, its first-nearest and its second-nearest neighbors). Thus, the interaction neighborhood $l$ consists of all nodes within the $l$th nearest neighbors of the node $i$.

### 2.3. The evolutionary algorithm

**Constructing a 2d Watts–Strogatz small-world network.** First, we build a regular square lattice with mean degree $\langle k \rangle = 4$ and

we then rewire some of its edges randomly with probability $p$. We thus follow the standard procedure for the generation of Watts–Strogatz small-world networks [65].

**Initialization.** At this stage, all properties are randomly assigned to each network site, i.e. one of the four strategies (ethnocentric, cosmopolitan, egoistic, or altruistic), one of the four tag colors, and one of the two cooperation preferences ('Yes' or 'No') are randomly selected and assigned to each site. Initially, each site is empty and owns the same baseline fitness $f_i = 0.12$. With respect to strategies, we employed two unconditional and two conditional strategies in our model. Altruism is one of the two unconditional strategies whereby agents always cooperate with everyone in the population (regardless of their tag colors), whereas egoism means that agents always defect (again, regardless of the opponents' tags). Agents using the ethnocentric strategy cooperate only with those opponents who share the identical tag color, whereas cosmopolitans cooperate only with the opponents having a different tag color; otherwise, they always defect. Thus, both ethnocentric and cosmopolitan agents are conditional strategists. Furthermore, we use a total of four distinct tag colors throughout the model simulations. The initial cooperation preferences of all ethnocentrics, cosmopolitans, and altruists are assumed to be other-regarding (i.e. 'Yes'), whereas the preferences of all egoists are initially set as self-interested (i.e. 'No').

**Evolutionary process.** There are six major evolutionary steps that constitute our main algorithm (at the beginning of each time step the list of living individuals is randomly shuffled):

1. **Immigration.** Starting with an initially unpopulated network, one agent per time step invades the system and lands on an empty, randomly selected site. The landing agent then inherits the properties of that occupied site, including the prescribed strategy, tag color, the cooperation preference, and the initial fitness.

2. **Interaction process.** In a random sequence order, each living individual interacts with the connected others within its designated interaction neighborhood. For example, in our basic model, an ethnocentric agent whose cooperation preference is other-regarding (i.e. 'Yes') will cooperate with its first living nearest neighbors (the maximum number of the first nearest neighbors is 4) if the tag color of these neighbors is the same as that of the ethnocentric agent. Similarly, cosmopolitan agents will cooperate if and only if their cooperation preferences are 'Yes' and their opponents with different tag colors are placed within their nearest neighborhood. In each cooperative interaction, a donor agent who cooperates pays a cost $c$ to donate a benefit $b$ to the recipient. Thus, each cooperative act is costly, whereas defection does not incur any costs. The benefits $b$ increase while the costs $c$ decrease an individual's fitness $f_i$.

3. **Conformity process.** Cooperation preference of each site $i$ is influenced by the cooperation preferences of the agents in the surrounding group $G_i$ focusing on $i$. The probability that the cooperation preference of $i$ is turned on or remains 'Yes' is then determined by the function shown in Eq. (1). The group $G_i$ contains all living individuals within $l$th nearest neighbors of $i$ (but does not contain $i$ itself). For example, since we study different interaction neighborhood sizes in the present paper, $l = 2$ means that group $G_i$ contains not only the first four nearest living neighbors but also the second nearest neighbors of agent $i$ (see e.g. Fig. 1).

4. **Reproduction.** Every living individual populating a site $i$ has a probability $f_i(t)$ that represents the fitness of an agent at a given time step $t$. This fitness is proportional to the probability of an agent's reproduction that can occur if at least one of its neighboring sites is empty. Thus, upon reproduction of $i$, the newborn baby agent will randomly occupy
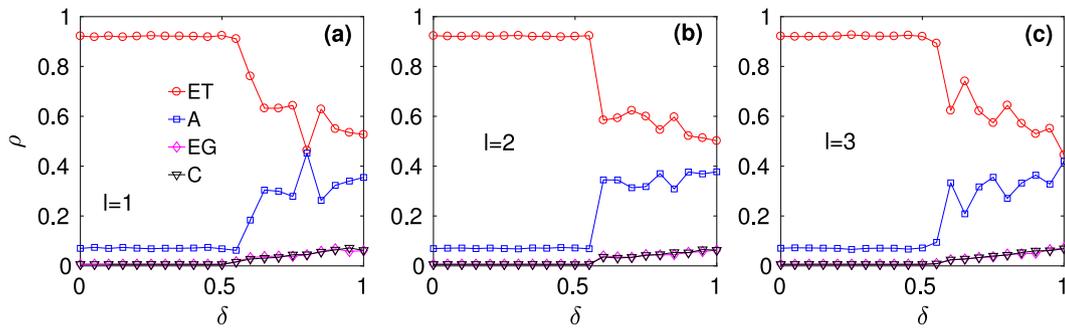
**Fig. 2.** The frequency $\rho$ of the four strategy types (ethnocentric (ET), altruistic (A), egoistic (EG), and cosmopolitan (C)) as a function of the strength of group influence $\delta$ for the tag-based cooperation model with conformity in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The displayed results are obtained by averaging over $N_r = 50$ independent realizations. The results in the three panels correspond to the three different interaction neighborhood sizes: $l = 1$ in panel (a), $l = 2$ in (b), and $l = 3$ in (c). The conversion sensitivity parameter was set to $\alpha = 10.0$, and the remaining parameter values were selected as described in the Methods section.

one of the neighboring empty sites of $i$, and then inherit the fitness and the cooperation preference of $i$. Moreover, with probability $\mu$, the baby will inherit a given strategy and tag color of parent $i$, but another strategy and another tag will be inherited with a probability $1 - \mu$.

5. **Death.** With probability $\nu$, any living individual in the system can be selected to die.

6. **Fitness renormalization.** Fitness $f_i$ of each individual is renormalized such that its average value is in agreement with the initial fitness value $f_i = 0.12$.

The above described evolutionary steps 1–6 are iterated until a stable equilibrium state has been attained. The averages over $N_r$ independent simulation realizations are then calculated to obtain the final results (the random number seeds were different for each simulation run). Moreover, a completely new small-world network with the same rewiring probability but with new links among network nodes was built for every tenth simulation run.

Unless otherwise specified, the baseline parameter values of our model were as follows: Mutation probability $\mu = 0.005$, death probability $\nu = 0.10$, cooperation cost $c = 0.01$, cooperation benefit $b = 0.03$, interaction neighborhood $l = 2$, network rewiring probability $p = 0.01$, and conversion sensitivity $\alpha = 10$. The strength of influence $\delta$ was always varied across the whole range of values $\delta \in [0\ 1.0)$. Finally, throughout the simulations, the size of the studied networks was $N = 200 \times 200$, meaning that the simulated systems had a total of 40,000 available network nodes.

## 3. Results and discussion

In the data analysis, we only included the simulation outcomes for which the system reached a stable, equilibrium state. For the analysis of the conditions in which the system exhibited oscillatory behavior, we only considered the time series obtained after sufficiently long evolutionary times for which the system did not show further significant changes of the observed oscillatory dynamics. For the simulated conditions reported herein, a stable equilibrium state was found mostly around $t = 10,000$, meaning that the results generated by our model at $t > 10,000$ were not significantly different from the outcomes observed at $t = 10,000$. In the following subsections, we separately present the results for the four different variants of our tag-based cooperation model with conformity.

### 3.1. Model I: Conformist altruists

In this model variant, the preferences of all types of strategists are under the influence of conformity, i.e., their cooperation preferences can be modified by the preferences of the surrounding

group members, influencing in turn their behavioral decisions (to cooperate or to defect). Egoists' preferences are also modifiable due to the group influence (i.e. they can change from 'No' to 'Yes' under the group influence), but these preference modifications cannot alter the actual behavior of egoists, because they always defect unconditionally. In addition, egoists' preferences can influence the preferences of other agents via the conformity rule. It is further assumed in this model variant that individuals only cooperate with their local neighbors; thus, no cooperative interactions occur via long-range connections. However, a cooperation preference of an individual can be influenced and modified via both local and long-range connections.

We can see in Fig. 2 that unlike the minority rule ($\delta < 0.5$), the application of the majority rule ($\delta > 0.5$) can facilitate the survival of altruists at various interaction neighborhood sizes $l$, even though the ethnocentric strategy dominates the population irrespective of $\delta$. The simulation result for the case without conformity is shown in Fig. 3(a), which nicely replicates a typical outcome of the standard tag-based cooperation model characterized by the strongly dominant ethnocentric strategy. Introducing the conformity mechanism into the model (Fig. 3(b)–(d)), we see that the overall level of ethnocentrism decreases, while the fraction of altruists increases with $\delta$. Notably, at $\delta = 0.6$, we observe an oscillatory coexistence of ethnocentrism and altruism in the model (Fig. 3(d)).

In Fig. 4, we show typical 2D color snapshots of the evolution of four strategies in the tag-based cooperation model with conformity. We can observe the relatively early onset of the ethnocentric superiority in the model at both weaker and intermediate strengths of the group influence $\delta$, with a transient dominance of altruism in the initial evolutionary stages $t \leq 300$ (see Fig. 4(a)–(b)). However, when the majority rule is in action at $\delta = 0.6$, after sufficiently long evolutionary time, the two dominant strategies (ethnocentrism and altruism) begin to coexist in the model (Fig. 4(c)) such that the ethnocentrism can no longer take over the population.

Fig. 5 further shows that relative to their competitors, especially in the initial evolutionary stages, altruists obtain much more help from other altruistic co-players, and as a result, their mean fitness is initially much higher than that of ethnocentric agents, irrespective of the strength of group influence $\delta$ (see Fig. 5(a)) and in spite of generally greater altruistic generosity towards ethnocentric individuals (Fig. 5(b)).

The situation then becomes more complicated as the number of agents populating the network grows and the clusters of altruists get in touch with ethnocentric ones, leading to a strong oscillatory behavior of the average fitness, particularly in later evolutionary stages (Fig. 5(a)). In the absence of the majority rule ($\delta \leq 0.5$), altruists located at cluster borders readily provide help to ethnocentric individuals while simultaneously not being timely
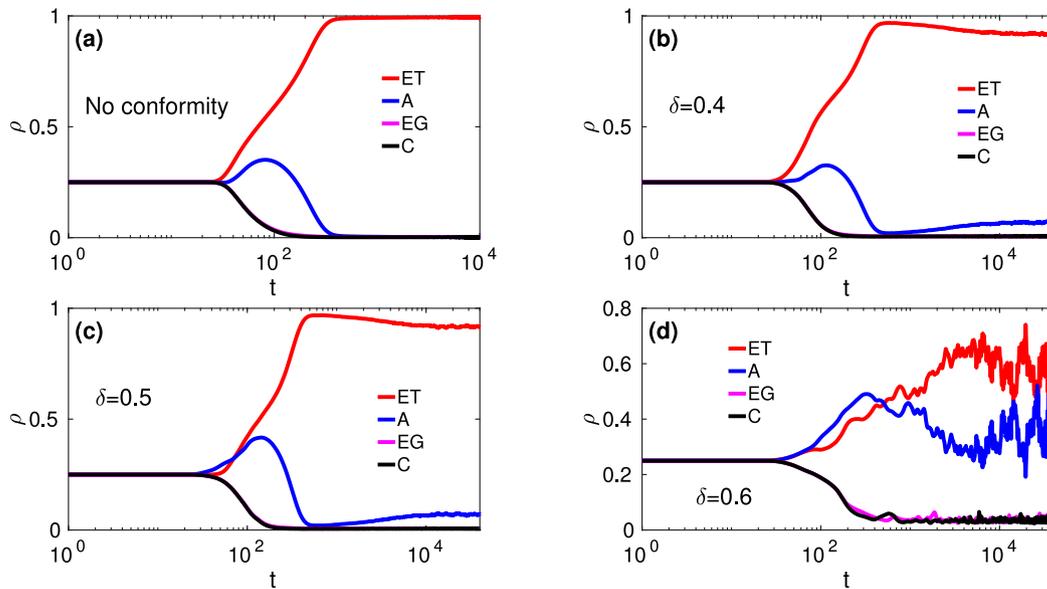
**Fig. 3.** The frequency $\rho$ of the four strategy types (ethnocentric (ET), altruistic (A), egoistic (EG), and cosmopolitan (C)) as a function of time in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The panel (a) shows the outcomes of the model simulations without conformity, while the remaining panels (b)–(d) correspond to the model with conformity under different values of $\delta$. In panel (b) $\delta = 0.4$, in (c) $\delta = 0.5$, and in (d) $\delta = 0.6$. In all four panels, $\alpha = 10.0$, $l = 2$, and the remaining parameter values were selected as described in the Methods section.
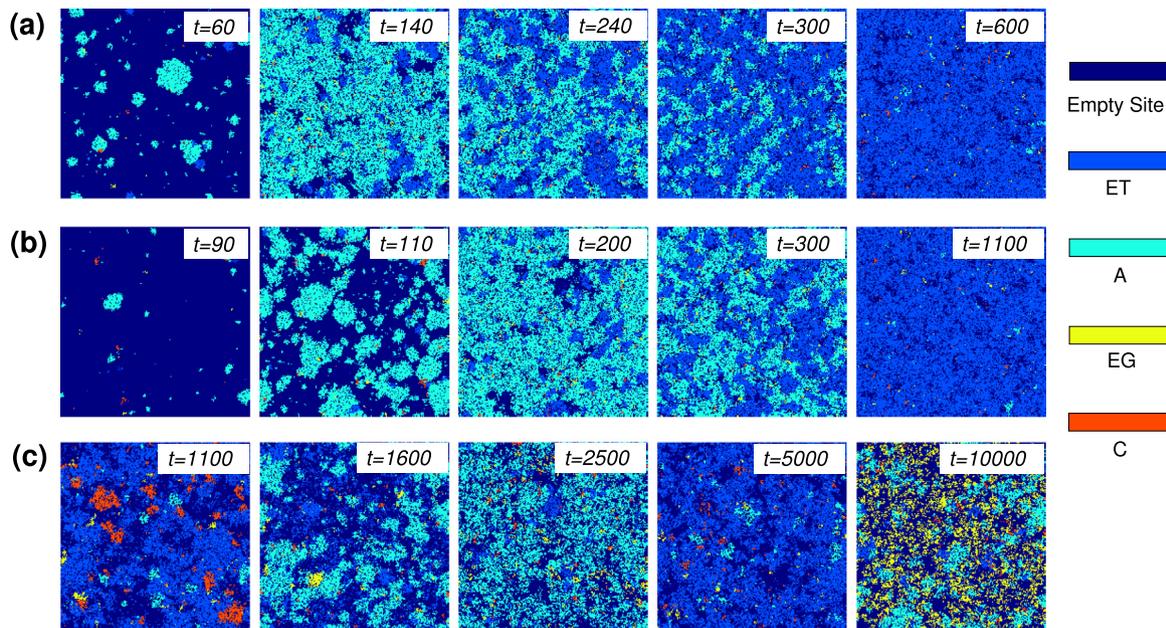


**Fig. 4.** Typical 2D color snapshots of the evolution of strategies in a tag-based cooperation model with conformity in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The minority rule was applied in simulations shown in panel (a) with $\delta = 0.4$ and in panel (b) with $\delta = 0.5$. The outcome of a simulation with the majority rule is shown in panel (c) with $\delta = 0.6$. The remaining parameters were taken as $\alpha = 10.0$, $l = 2$, and for the rest of parameters the baseline values were used as described in the Methods section.

compensated by other altruists or ethnocentrics (see Fig. 4(a)(b) and Fig. 5(b)). Due to the influence of the surrounding ethnocentric agents, the other-regarding preferences of altruists become reduced with time (see Fig. 5(c)). The borders of altruistic clusters then naturally begin to shrink back, and the ethnocentric clusters gradually expand at the expense of the altruistic domains (see Fig. 4(a)(b)).

After a sufficiently extended evolutionary time, the whole network becomes nearly fully saturated with ethnocentric individuals for $\delta \leq 0.5$. As the majority rule begins to work with $\delta > 0.5$, the altruists at their cluster borders become less likely to provide help to the surrounding ethnocentric individuals. The altruists then

receive more help from other altruists thereby stabilizing their own clusters (see Fig. 4(c)). However, when the periods in which altruists assist their own kind become no longer more frequent than the periods in which ethnocentric agents help their tag-mates (but instead they both begin to oscillate in evolutionary time as shown in Fig. 5(d)), the two groups of strategists (altruists and ethnocentrics) start to dynamically coexist and can no longer outcompete each other (see again Fig. 3(d)).

Since the other two groups of strategists in our model (egoists and cosmopolitans) are strongly suppressed, as has been the case in most previous models of tag-based cooperation, we focus our further analyses mainly on the dynamical behaviors of altruistic
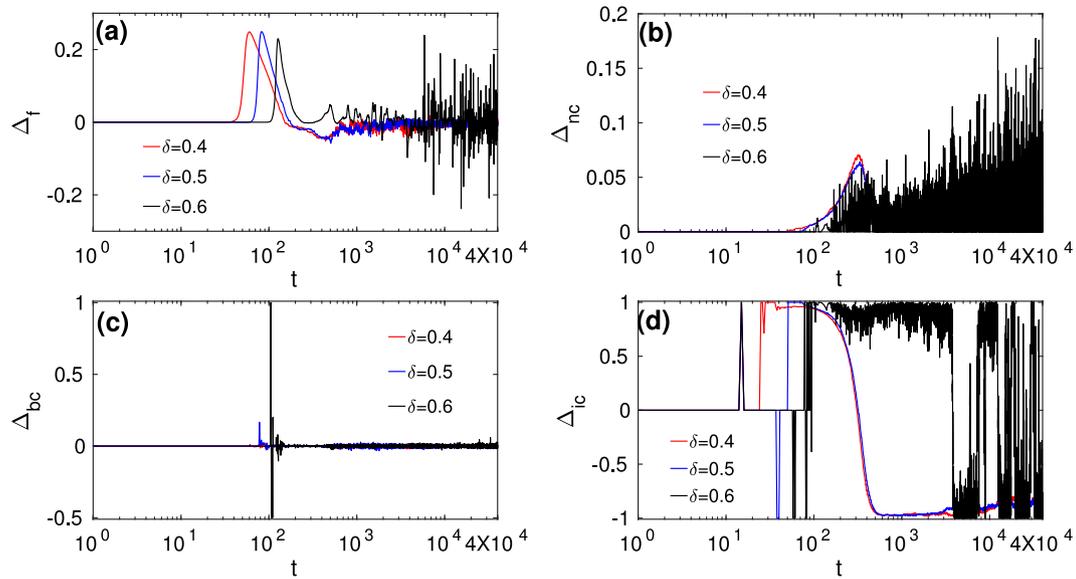
**Fig. 5.** (a) Evolution of the difference between the average fitness values of altruists and ethnocentrics $\Delta_f(t) = \overline{f_A}(t) - \overline{f_{ET}}(t)$. Positive $\Delta_f(t)$ values indicate that altruists own higher fitness than ethnocentric agents; the peaks around $t = 10^2$ correspond to the rapid expansion of altruists before their contact with ethnocentric clusters occurs. (b) Evolution of the difference between the frequency of cooperative acts provided by altruists to their connected ethnocentric opponents $n_{AE}(t)$ and cooperative acts provided by ethnocentric agents to their connected altruistic opponents $n_{EA}(t)$, i. e., $\Delta_{nc}(t) = (n_{AE}(t) - n_{EA}(t))/n_{coop}(t)$, where $n_{coop}(t)$ represents the total number of cooperative acts in the system at time $t$. Higher $\Delta_{nc}(t)$ indicates more cooperative acts by altruists provided to their ethnocentric neighbors (than vice versa), which reveals that altruists significantly contribute to the growth and increase of ethnocentric clusters. (c) Evolution of the difference between the average cooperation preference of altruists connected to ethnocentric agents and the average cooperation preference of ethnocentric agents connected to altruists $\Delta_{bc}(t)$. The positive peak of $\Delta_{bc}(t)$ indicates a higher cooperation preference of altruists than of their ethnocentric neighbors, i. e., the overall greater generosity of altruists which is ultimately leading to a decay of their average fitness. (d) Evolution of the difference between the frequency of cooperative acts exchanged among altruists and the frequency of cooperation exchanged among ethnocentric agents $\Delta_{ic}(t)$. $\Delta_{ic}(t) \simeq 1.0$ indicates that altruists obtain sufficient support from other altruists within their own clusters helping them to successfully resist the invasion of ethnocentric intruders. The results shown in all four panels correspond to simulations conducted on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The remaining parameters were taken as $\alpha = 10.0$, $l = 2$, and for the rest of the parameters their baseline values were selected as described in the Methods section.

and ethnocentric agents. Nevertheless, relative to the condition without conformity or with the influence of the conformity mechanism but under the minority rule, we find that the majority rule ($\delta > 0.5$) in our model can actually give rise to visibly more egoists and cosmopolitans in the population (see Fig. 4(c)).

### 3.2. Model II: Nonconformist free altruists

In this version of our model, the preferences of altruistic agents are not conformity biased, i.e., their cooperation preferences cannot be modified by the preferences of the surrounding group members, and as a result, their behavioral decisions cannot be affected. When all altruistic agents are spared from the influence of conformity, as is the case in the current model variant, we then call them the *free altruists*. In other words, irrespective of the preferences of other surrounding agents, the free altruists carry the cooperative preference 'Yes' throughout the simulation, and consequentially, they always cooperate unconditionally with all interacting neighbors. At the same time, the preferences of ethnocentric, cosmopolitan, and egoistic agents are under the influence of conformity in this model version. However, as in the previous model variant, although egoists' preferences remain modifiable due to the group influence (i.e. they can change from 'No' to 'Yes' under the group influence), this conformity bias cannot alter the actual behavior of egoists because they always defect unconditionally. Moreover, egoists' preferences can influence the preferences of others via the conformity rule. It is further assumed, as in the previous model variant, that individuals only cooperate with their local neighbors; thus, no cooperative interactions occur via long-range connections. However, a cooperation preference of an individual can be influenced via both local and long-range connections. In Fig. 6 we see that in the tag-based cooperation model with conformity and free altruists, the majority rule ($\delta > 0.5$) stably promotes the

dominance of altruism. The advantage of altruists in this model version remains robust against the variation of the conversion sensitivity $\alpha$ and the size of the interaction neighborhood $l$ (see again Fig. 6).

In Fig. 7, we show the evolutionary dynamics of the four strategy types for the three different strengths of group influence $\delta$ in the conformity model with nonconformist free altruists (Fig. 7(b)–(d)), and the basic tag-based cooperation model without conformity (Fig. 7(a)). Relative to the results obtained with the first model variant (where all agents are influenced by the conformity mechanism, shown in Fig. 3), we observe a remarkable outbreak of altruism in the model with nonconformist free altruists, in which altruistic agents start to dominate the population already at the threshold value of $\delta = 0.5$ (Fig. 7(c)).

Comparing the 2D color snapshots of the spatial strategy distributions in Figs. 4 and 8, we can see that when altruists are not susceptible to conformity (Fig. 8) while the remaining agent types conform to the influence of others, altruists start to produce more dense and larger altruistic domains that can persist in evolutionary time, thereby preventing the formation of compact ethnocentric clusters. This phenomenon is observed already at the threshold value of the group influence $\delta = 0.5$, and becomes even stronger pronounced under the majority rule $\delta > 0.5$ (see Fig. 8(b)–(c)).

We identify three potential mechanisms that are relevant to the dominance of altruism under the majority rule ($\delta > 0.5$) in the tag-based model with nonconformist free altruists: (1) When the majority rule is in action ($\delta = 0.6$), free altruists can significantly reduce their overall generosity towards the surrounding ethnocentric agents (see Fig. 9(b)), thereby maintaining higher average fitness than under $\delta < 0.6$ (see Fig. 9(a)). (2) Concomitantly, given the limitations of the available identical tag colors and the continuous modification of cooperation preferences (due to conformity), the intra-tag ethnocentric cooperation declines and the ethnocentric
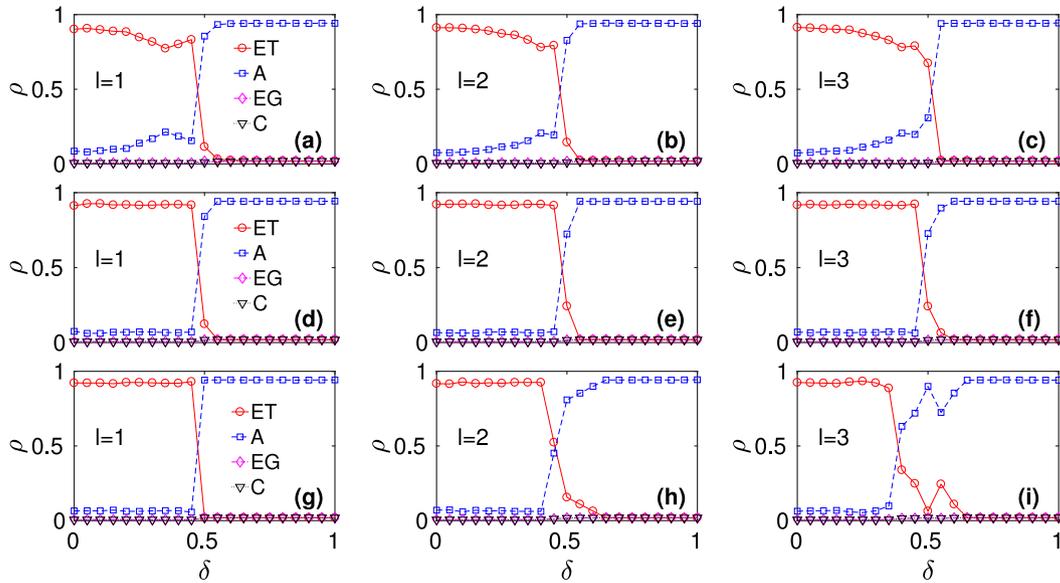
**Fig. 6.** The frequency $\rho$ of the four strategy types as a function of $\delta$ for the model version with nonconformist free altruists evolving on a 2d WS small-world network with periodic boundary conditions, the rewiring probability $p = 0.01$, the varying interaction neighborhood sizes $l$, and different conversion sensitivities $\alpha$. The condition with the interaction neighborhood $l = 1$ is shown in the left-column panels, $l = 2$ in the middle-column panels, and $l = 3$ in the right-column panels. The conversion sensitivity $\alpha = 5.0$ condition is shown in the top row of panels (a–c), $\alpha = 10.0$ in the middle-row panels (d–f), and $\alpha = 20.0$ in the bottom-row panels (g–i). All results represent the averages taken over $N_r = 50$ independent realizations.
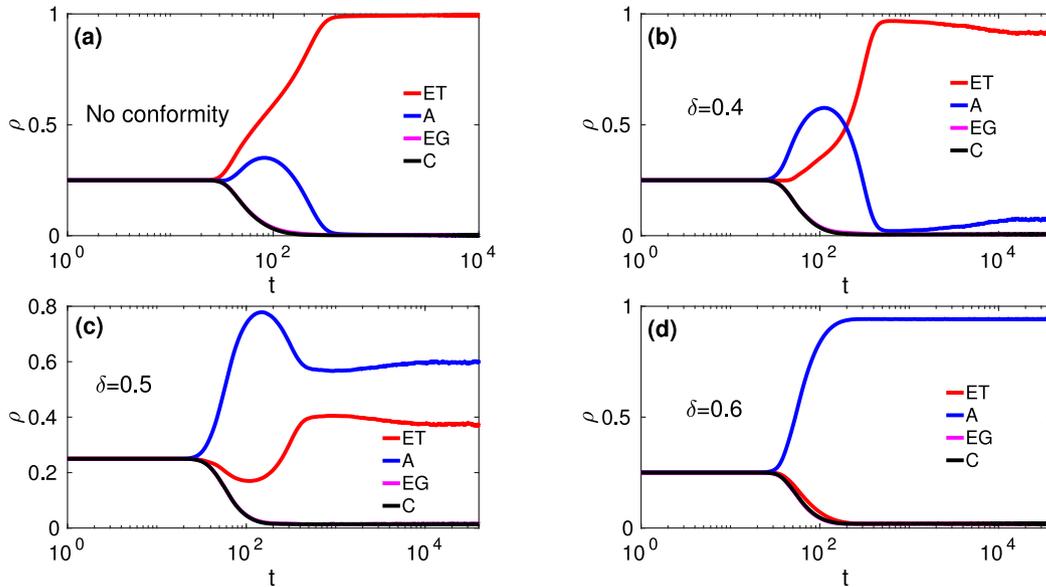


**Fig. 7.** Evolution of the four strategy types in a tag-based cooperation model with nonconformist free altruists and conformity biased transmission of other strategies in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. In (a), the baseline tag-based cooperation model without conformity is shown. In (a)–(c), the tag-based cooperation model with the added conformity component (but with nonconformist free altruists) is displayed. Furthermore, the panel (b) corresponds to the case with $\delta = 0.4$, panel (c) with $\delta = 0.5$, and the outcome for the system simulated under the majority rule $\delta = 0.6$ is shown in panel (d). The conversion sensitivity for the cases with the conformity mechanism was taken as $\alpha = 10.0$ and the interaction neighborhood size was $l = 2$. The values of the remaining parameters were taken as in the baseline condition, as described in the Methods section.

agents do not attain a sufficiently high average fitness in order to reproduce as quickly as altruists (see again Fig. 9(a), where the persisting positive value of $\Delta_f(t)$ under $\delta = 0.6$ clearly shows that altruists own higher fitness than ethnocentric agents). (3) Consequentially, under the majority rule and in the long-term, the generosity remains better preserved within altruistic than within ethnocentric clusters (as evidenced in Fig. 9(d)).

In addition, for this model variant with nonconformist free altruists and only local interactions among connected agents, we also examined the role that the varying long-range connections may play in the system evolution (even though only preference

modifications but not cooperative acts can occur along them). In Fig. A.16 (see Appendix A), we can see that the number of long-range connections (with the rewiring probability $p$ ranging from $p = 0.01$ to $p = 0.3$) does not significantly affect the simulation outcomes when the interactions among individuals are still strictly local.

### 3.3. Model III: Pinning free altruists

We have seen thus far that pure altruism can outcompete other strategies and stably dominate the system when altruistic agents
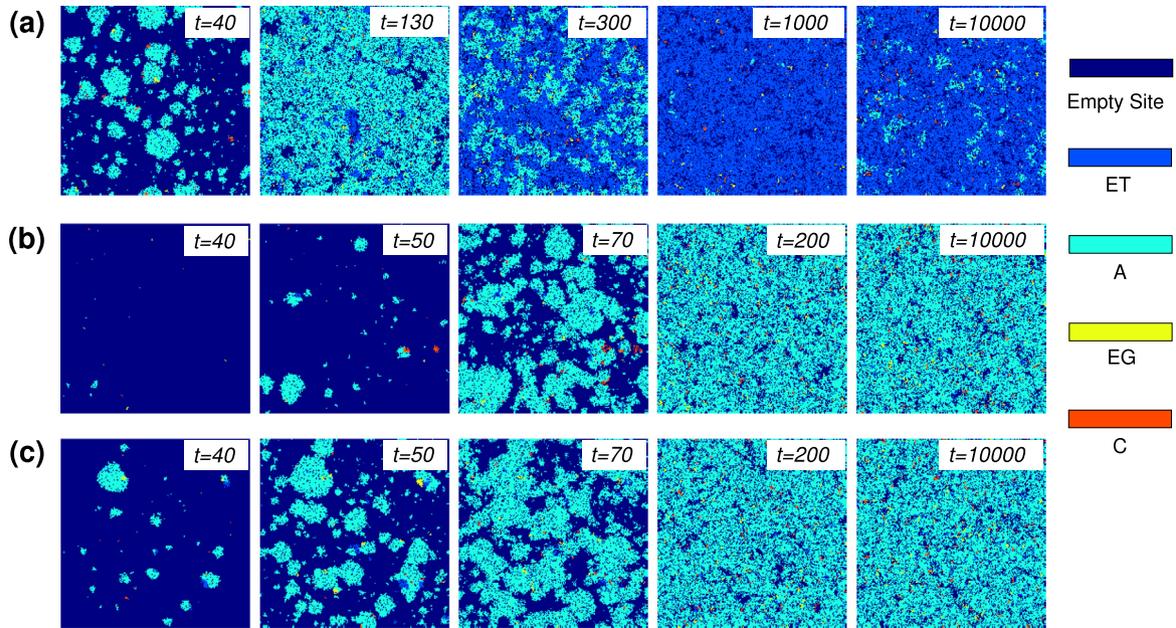
**Fig. 8.** Typical 2D color snapshots of the evolution of four strategies on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$ for the model version with conformity and nonconformist free altruists. In panel (a) $\delta = 0.4$, in (b) $\delta = 0.5$, and $\delta = 0.6$ in panel (c). The remaining parameters were taken as $\alpha = 10.0$, $l = 2$, and the rest as in the baseline model, as described in the Methods section.
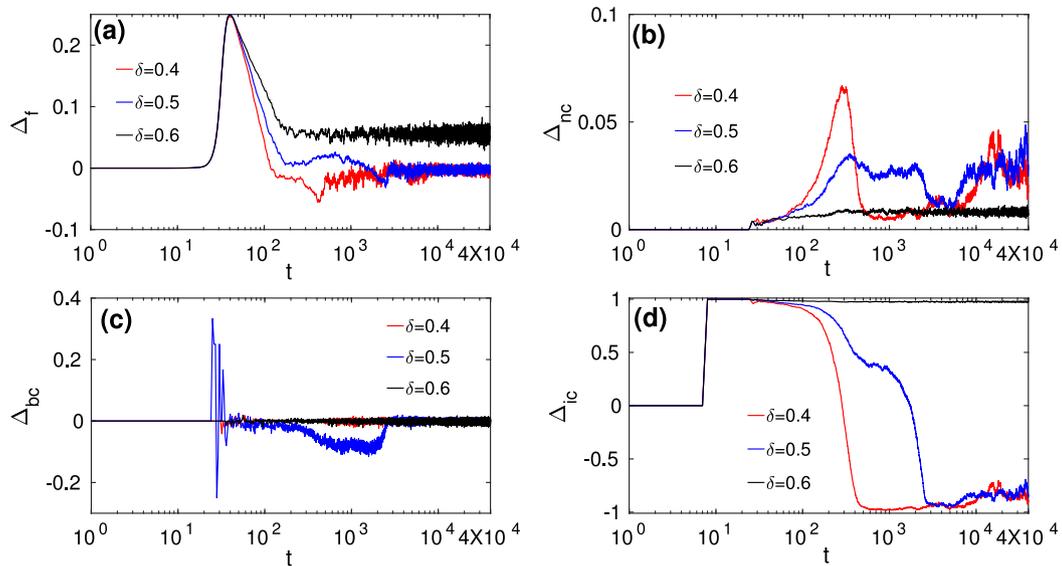


**Fig. 9.** Dynamical behavior of the average fitness $\Delta_f(t)$ (a), the evolution of $\Delta_{nc}(t)$ (b), of $\Delta_{bc}(t)$ (c), and of $\Delta_{ic}(t)$ (d) for the three different values of the group influence $\delta$. All outcomes correspond to simulations on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$ with the conformity mechanisms but with nonconformist free altruists. The remaining parameter values were taken as $\alpha = 10.0$, $l = 2$, and for the rest, as described in the Methods section.

do not conform, while all other strategists are conformity biased. Since in the previous model version all altruists were nonconformists, we now investigate the question of what is the minimal number of nonconformist altruists required to drive the tag-based system into the desired state of pure altruism. To this effect, we introduce and test a mechanism for the systematic control of the evolutionary dynamics of altruism in a tag-based model with conformity.

In evolutionary game theory, a widely used method to control the dynamics of prosocial behavior, i.e. to increase and sustain the frequency of cooperators and concomitantly to decrease the level of defectors, is to seed pinning individuals such as super altruists, who are forced to never change their strategy throughout the system evolution [66,67]; for a related application of this

method to opinion dynamics, see Ref. [68]. Inspired by this earlier research, in the present model version we study the consequences of introducing a random pinning control mechanism in the tag-based cooperation model with conformity, in which the varying numbers of pinning free altruists are tested for their resistance to the group influence. Naturally, the associated effect of this control mechanism on the evolutionary dynamics of cooperation in a tag-based system with conformity is systematically studied.

In our control mechanism, we simply let a certain randomly selected proportion $f_p$ of altruists to be the free altruists. Unlike other altruists, the pinning free altruists will not be influenced by the cooperation preferences of others, i.e., their cooperation preference will remain 'Yes' throughout the system evolution. Moreover, it is assumed in this model variant, as in all previous ones, that the
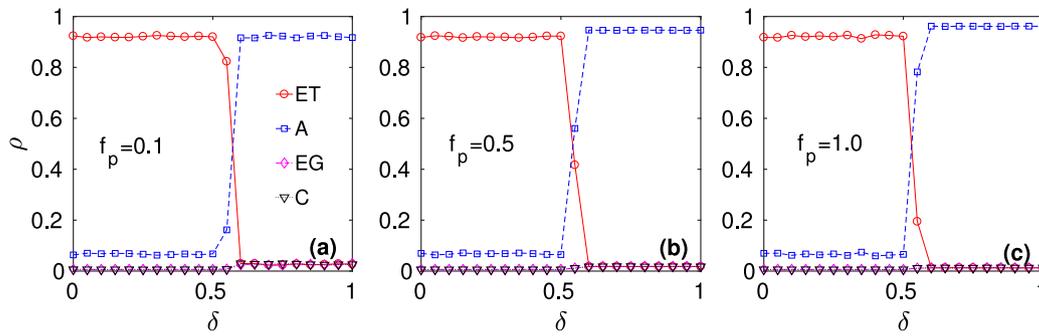
**Fig. 10.** The frequency $\rho$ of the four strategies as a function of $\delta$ for different degrees of the pinning control in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$; $f_p$ represents the fraction of nonconformist altruists, i.e. $f_p = 0.1$ in (a), $f_p = 0.5$ in (b), and $f_p = 1.0$ in (c). The shown results are the averages over $N_r = 50$ independent realizations. The remaining parameters were $\alpha = 10.0$, $l = 2$, and for the rest as described in Methods.
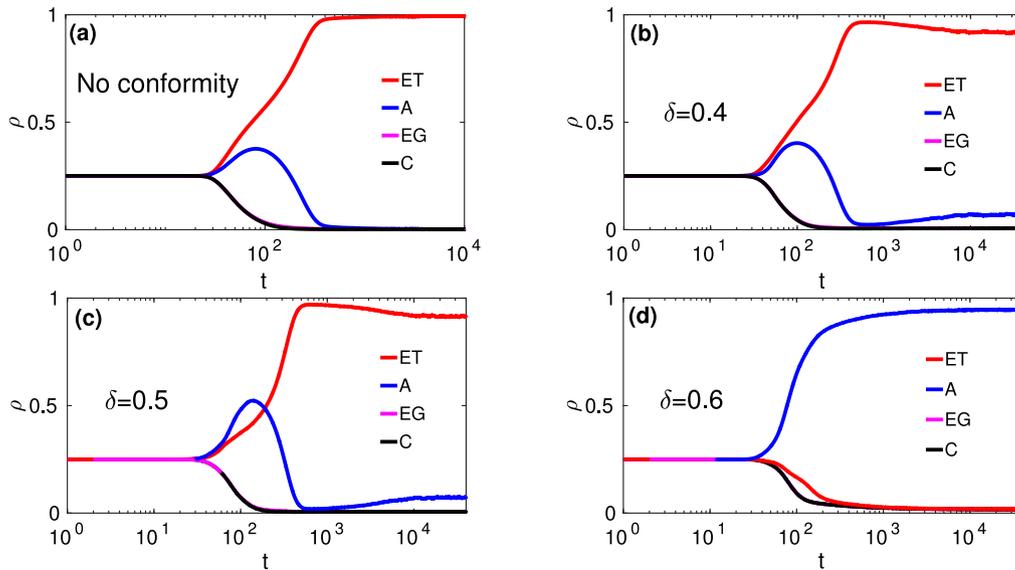


**Fig. 11.** Evolution of the four strategy types on a 2d WS small-world network with periodic boundary conditions, the rewiring probability $p = 0.01$, and with nonconformist altruists. In (a) conformity is not introduced. In (b)–(d), conformity is introduced. Furthermore, (b) $\delta = 0.4$, (c) $\delta = 0.5$ and $\delta = 0.6$ in (d). The other parameters were $\alpha = 10.0$, $l = 2$, and the rest as described in Methods.

cooperative interactions can only occur across local connections in the network, whereas individual preferences can be biased both via local and global (long-range) connections.

Fig. 10 demonstrates that under the majority rule ($\delta > 0.5$) and with the applied pinning control mechanism, not all altruists have to be free from conformity bias in order to promote the dominance of pure altruism in the model. Instead, it is sufficient for the stable altruistic dominance that only a small fraction of altruistic agents such as $f_p = 0.1$ is not conformity biased, since the effect is not significantly enhanced if the degree of pinning altruists $f_p$ is further increased. We also see in Fig. 10 that the pinning altruists' influence is offset when the individual preferences are easily biased by the preferences of only a few surrounding agents, i.e., when the minority rule ($\delta < 0.5$) is in action, irrespective of the actual degree of the pinning control in the model (see Fig. 10(b)(c)).

The evolution of the system for the cases without conformity and with the conformity bias for different values of $\delta$ and $f_p = 0.1$ is shown in Fig. 11. Here, we see again that the majority rule is indeed necessary to promote the altruistic dominance in the population, which cannot occur under the minority rule. This is also evidenced by the spatial snapshots shown in Fig. 12, which additionally demonstrate that under the minority rule and the implemented pinning control mechanism with $f_p = 0.1$, the altruistic clusters are prevented from expansion rather early in evolutionary time,

relative to the previous model version in which all altruists are free from the conformity bias.

Conformity biased altruistic agents located near the pinning altruists can get enough help to maintain positive average fitness at $\delta = 0.6$ (see Fig. 13). In the presence of the majority rule, clusters with pinning altruists can avoid cooperation with ethnocentric agents and better preserve their in-group generosity than ethnocentrics (see Fig. 13(b) and (d)). However, we can also see in Fig. 13 that the asynchrony in the cooperation preferences between the pinning free altruists and conformity biased altruists is causing a stronger oscillatory behavior relative to the previously studied model version in which all altruists are nonconformists. Additionally, compared with the cases shown in Figs. 5 and 9, the average fitness of altruists is much lower under the pinning control mechanism (see Figs. 13 and A.19).

Remarkably, in the presence of the majority rule ($\delta > 0.5$), even a very small number $f_p$ of pinning altruists can push the system into the phase of nearly full-altruists, irrespective of the underlying network rewiring probability (see Fig. A.14). We thus see in Appendix A in Figs. A.14 and A.17–A.19, that seeding initially only a tiny fraction of pinning free altruists (but at least $> 1\%$) can establish and maintain the other-regarding cooperation preferences of most altruists, thereby greatly facilitating their superiority throughout the system evolution. There exists a saturation threshold (order of magnitude of $10^{-2}$) beyond which the system
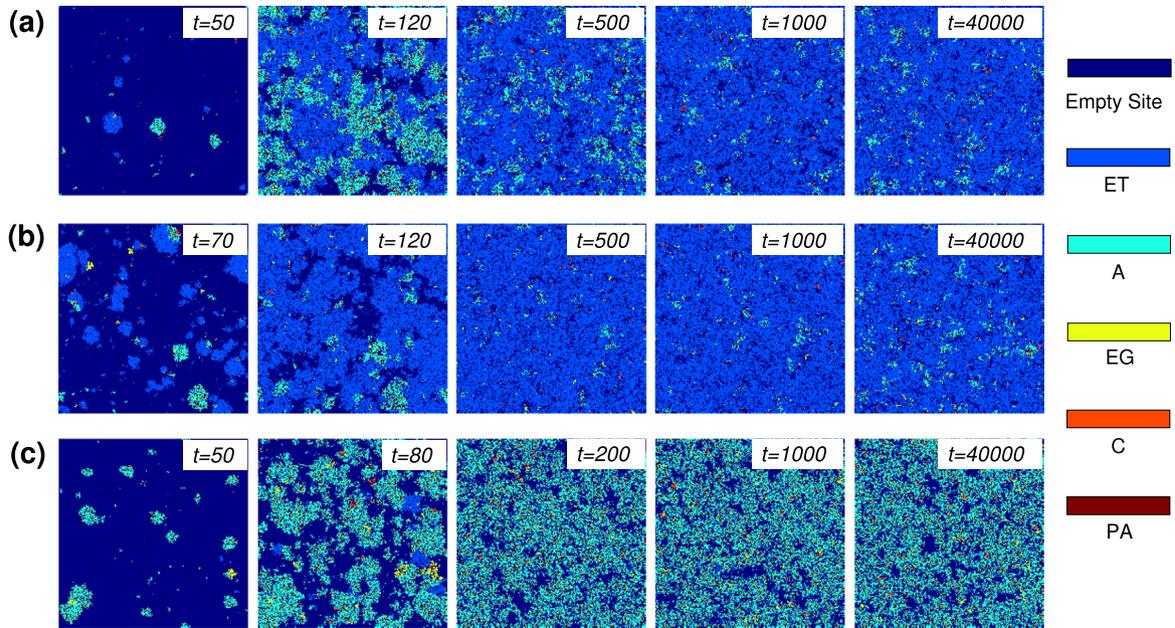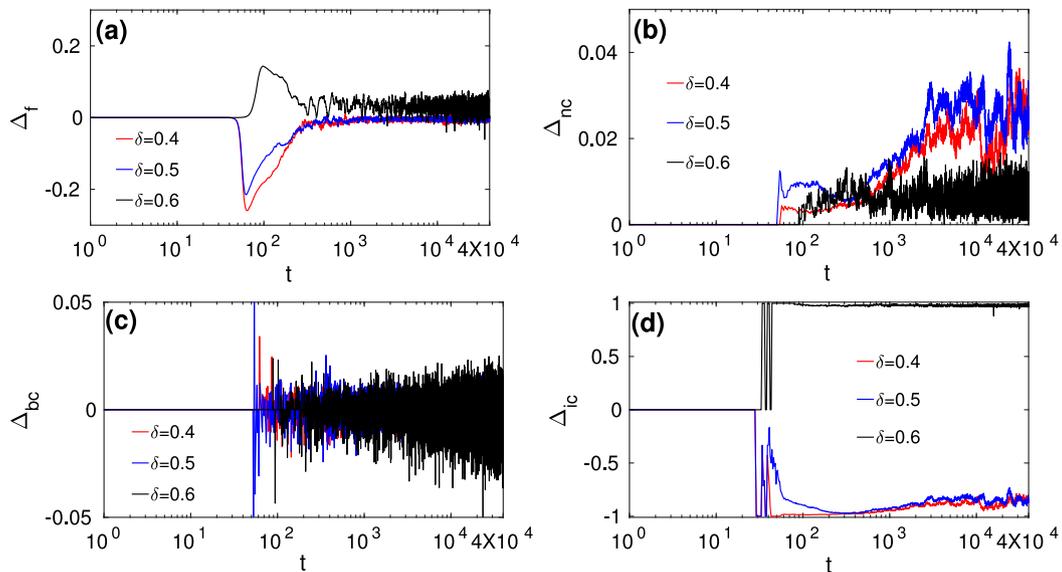
**Fig. 12.** Typical 2D spatial snapshots of the evolution of four strategies on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$, where conformity influences the cooperation preferences of ethnocentric, cosmopolitan and altruistic agents but not the preferences of pinning altruists. Initially, $f_p = 0.1$ of altruists are randomly set as free (nonconformist) altruists. Furthermore, (a) $\delta = 0.4$, (b) $\delta = 0.5$ and $\delta = 0.6$ in (c). The other parameters were $\alpha = 10.0, l = 2$, and for the rest as described in Methods.



**Fig. 13.** (a) Dynamical behaviors of $\Delta_f(t)$ for three different values of $\delta$. (b) Dynamical behaviors of $\Delta_{nc}(t)$ for three different values of $\delta$. (c) Dynamical behaviors of $\Delta_{bc}(t)$ for three different values of $\delta$. (d) Dynamical behaviors of $\Delta_{ic}(t)$ for three different values of $\delta$. The system evolution was simulated on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. Initially, $f_p = 0.1$ of altruists were randomly set as free, nonconformist altruists. The other parameters were $\alpha = 10.0, l = 2$, and the rest as described in Methods.

is nearly full of altruists, such that further pinning control becomes redundant for driving the direction of the system evolution.

### 3.4. Model IV: Effects of long-range connections

In this model version, we explore the system evolution for the case that individuals can interact with each other not only via local links but also along global, long-range network connections. Thus, in this model variant, both cooperative acts and preference biases occur via both local and global connections in the system.

The simulations results shown in Appendix A in Figs. A.15 and A.20–A.22, are based on the model where all altruists are initially free from conformity and where both cooperative interactions

and the preference bias occur through both short and long-range connections. Surprisingly, we can observe that the overall pattern of results does not change significantly when interactions occur also via global, long-range network connections. However, we can observe a modest downfall of altruists with an increasing network rewiring probability $p$ when interactions occur also along the global network links.

In Fig. A.23 we show the condition with conformist altruists and interactions via both local and global network connections. Thus, in this condition, preferences of all agent types (including altruists) are conformity biased and influenced by the preferences of other group members. It can be observed that this restriction indeed brings altruists into a disadvantaged position, since they can no

longer dominate the population irrespective of $\delta$. Moreover, we see again that increasing the long-range connections via stronger rewiring of the network further reduces the number of altruists in the population. Yet the unfavorable conditions with more long-range connections do not qualitatively change the final evolutionary outcomes.

Finally, detailed results of our mathematical analyses are given in Appendix B. Our main analytic results can be summarized as follows: We first replicated the typical finding showing that in the absence of any mechanisms for the evolution of cooperation, there is the dominance of egoism in the model. Together with our computational experiments, we further concluded that when all types of strategists are susceptible to conformity, neither conformity bias nor the spatial structure alone can tune the evolution in favor of altruists. However, when compared to previous computational experiments with structured populations, our further analysis and numerical simulations of the aspatial model revealed that in combination, spatial structure and nonconformity of cooperators can give rise to a stable dominance of altruists in tag-based multiagent systems.

## 4. General discussion and outlook

### 4.1. Conforming vs. nonconforming altruists

In the absence of conditional strategies, previous studies have shown that conformity may support cooperation but often only under specific conditions or in the presence of other cooperation-promoting mechanisms (e.g. [37,44]). For example, recent studies have revealed that the effects of conformity on cooperation are highly dependent upon the choice of the underlying network topology [46], such that in certain interaction structures (e.g. scale-free networks), conformity can even hinder the emergence of cooperative behavior. Thus, even in the presence of only simple unconditional strategies, previous studies addressing the influence of conformist bias on cooperation have remained largely inconclusive, suggesting that conformity alone cannot explain the cooperation levels typically observed in social systems [43].

In our present study, we have addressed for the first time the combined effects of conformity and tag-mediated interactions, in which both conditional and unconditional strategies are involved. We found that spatial structure and nonconformity of unconditional cooperators are the two indispensable ingredients for the stable persistence of dominant pure altruism in tag-based multiagent systems, thereby corroborating the previously detected essential role of interaction structure in determining conformity's cooperation-promoting potency [46]. However, to our knowledge, none of the previous studies has reported on the relevance of nonconforming unconditional cooperators in combination with conformity-biased transmission of other strategies in tag-based multi-agent networked systems. In this context, our study has revealed a series of non-trivial findings which we henceforth individually discuss.

Although the ethnocentric strategy remains dominant after introducing the conformity into the model (Fig. 3(b)–(d)), the overall level of ethnocentric cooperation decreases, while the fraction of altruists increases with $\delta$. Notably, at $\delta = 0.6$, we observed an oscillatory coexistence of ethnocentrism and altruism (Fig. 3(d)). Interestingly, a previous study has shown that such oscillatory behavior can originate from the synchronization of individual strategies due to the locally introduced conformist bias [69]. We have shown, however, that in a tag-based cooperation model with multiple strategies, the oscillatory behavior can emerge if conformity bias spreads both locally and globally over the long-range connections in the network.

In the model variant with conformity bias but with free non-conforming altruists, we have found that the majority rule ($\delta >$ 0.5) stably promotes the outbreak of the dominant altruism. The advantage of altruists in this model version remains robust against the variation of the conversion sensitivity $\alpha$ and the size of the interaction neighborhood $l$. It is thus indeed possible that pure altruism can emerge and outcompete other strategies in the system when altruistic agents do not conform to the group, while at the same time, all other strategists are conformity biased.

The vast majority of earlier studies on tag-based cooperation (e.g. [10,12,24]) reported stable dominance of the intra-group 'ethnocentric' strategy, that consistently outweighs all other competing strategies across a number of conditions. More recently, only a few studies [17–19,23] have identified conditions under which non-ethnocentric strategies can also dominate the tag-based system, thereby reliably suppressing in-group biased or other selfish strategies. To our knowledge, our present study is the first one to identify a novel condition, this time within the framework of a co-evolutionary model, under which pure altruism, i.e. unconditional cooperation, can globally outcompete tag-based conditional strategies and unconditional defection.

Using a random pinning control mechanism, we found that only a small fraction of nonconformist altruists (but at least $>$ 1%) in the system can establish and maintain the other-regarding cooperation preferences of most altruists, thereby greatly facilitating their superiority throughout the system evolution. For practical considerations, this result suggests that controlling only a subset of altruists, the state of a nearly complete altruism in the population can be attained while simultaneously keeping the control costs at a minimum.

We have thus shown that a stable control of the system in favor of altruists is possible even without the application of degree-based [66] or other more sophisticated control mechanisms [67]. Instead, a simple random selection of a small fraction of pinning free altruists is sufficient to drive the system towards the globally dominant state of pure altruism. Excitingly, we have checked that this finding is highly robust and independent of the number of long-range connections in the network, the size of the interaction neighborhood, or the conversion sensitivity $\alpha$. Of course, larger degree of control would further increase the robustness of altruism and its resistance to the influence of other types of strategies, and moreover, it would speed up the emergence of an altruistic equilibrium in evolutionary time.

When cooperative acts and conformity bias are allowed to occur via both local and global connections in the system, we observed no significant changes relative to the previously observed result patterns, though the rewired interactions do tend to cause a slight decrease in the overall level of altruism.

Our results are in line with a more recent study [43] suggesting that the co-evolution of cooperation and conformity alone cannot explain the observed levels of cooperation typically found in human societies, but that instead conformity may play a role when combined with other relevant factors. Specifically, our computational experiments, the mathematical analysis, and additional numerical simulations revealed that under the majority rule, spatial structure and nonconformity of cooperators in combination are the two indispensable ingredients for the stable dominance of altruism in tag-based multiagent systems.

In agreement with previous findings [16,21], we have seen that the aspatial tag-based cooperation model, without any population structure, can indeed give rise to elevated intra-tag ethnocentric cooperation. However, the same aspatial model, even in the presence of the majority rule, was not able to promote global altruism to the extent that it outcompeted all other strategies (see e.g. Fig. B.27 in the Appendix). Instead, we found that this was only possible when the spatial structure was combined with nonconforming unconditional cooperators, while concomitantly, the transmission of other strategies was conformity biased (cf.

Figs. 6, B.27 and B.31). Our analytic arguments and multiagent networked simulations have further revealed that spatial structure is a necessary but not a sufficient condition for the emergence of global altruism in tag-based cooperative systems.

In the context of real-world social phenomena, our model provides an alternative, nonlocalist explanation for the emergence of tag-based cooperation and the associated coexistence of ethnocentrism and altruism that is ubiquitous in human and other animal societies. In its standard localist view, it has been assumed that preferential cooperation with group members and the resulting global ethnocentrism are largely associated with the perceived shared similarity among locally interacting tag-mates. However, our findings suggest that such local processes of tag detection, tag-mediated interaction among neighboring agents, and the local conformity-driven influence of strategies under the minority rule, can only explain the emergence of ethnocentric dominance in tag-based cooperation models.

Global processes, on the other hand, such as the pressure to conform to a collectively desirable group norm under the majority rule dictate alternative outcomes in tag-based systems, characterized by either oscillatory coexistence of ethnocentric and altruistic agents, or by the equilibria with dominant nonconforming altruists. In agreement with a recent observation [22], our model thus suggests that ethnocentrism is not that easily evolvable and as robust as has been previously assumed, since ethnocentric clusters can become vulnerable to invasion by pure altruists in the presence of influential social norms.

It has been suggested previously [50] that altruistic norm compliance is shaped by tag-mediated ingroup favoritism. To our knowledge, our current model represents the first explicit implementation of this view where normative obedience of agents depends on the group membership of their opponents. For example, if agents are ethnocentric intra-tag cooperators, their cooperative decision making will be affected in our model by the conformity bias if and only if their interacting opponents share the same tags. We have demonstrated that one consequence of this dependency of norm compliance on group membership in tag-based systems is a possible dominance of global pure altruism, provided that spatially structured unconditional cooperators refuse to obey the enforced social norm. Our results thus suggest that this dependency of norm obedience on group membership may have evolved not just to promote cooperation within groups, but also to establish global indiscriminate cooperation in the society, especially in the presence of strong ethnocentric individuals and selfish unconditional defectors.

### 4.2. Future research directions

Earlier studies have evidenced that conformity interacts not only with the type of the game that is being played among agents [44,49], but also with the type of the underlying network topology [46,70]. Since our present model simulations were conducted on small-world networks with varying rewiring probability [65, 71], we suggest that next extensions of our work should investigate the dynamics of tag-based cooperation with conformity in other network topologies such as scale-free networks. Importantly, studying the interactions between multiple groups and the conformity biased transmission of cooperative strategies over different layers of multiplex networks would introduce a further level of complexity [49], that rightfully deserves special attention in future studies.

In this context, one possible generalization of our model could consist of a case where agents play one game with their ingroup mates but another game against the outgroups, which would in turn enable a more systematic study of the intergroup bias [72] and its potential combined effects with conformity on tag-mediated

cooperative behavior. In addition, the potency of the conformity mechanism in combination with punishment [28,71,73] and reward incentives [23,74] should be elucidated in future studies, especially in the context of systems with high diversity whose collective intelligence often depends upon non-trivial, optimally designed incentive schemes [23,75].

### 4.3. Potential application areas

An often overlooked problem with tag-based systems is that tag-mediated cooperation remains largely preserved within the tag-sharing ingroup clusters, which in the long-term, can endanger the diversity of the population and lead to the associated productivity gaps or innovation shocks. It may therefore be useful to develop tag-based cooperation systems that are not only dominated by intra-group cooperation but which can instead also lead to global cooperative outcomes with dominant unconditional cooperation. In particular, such outcomes based upon hybrid tag-based systems with conformity may be desirable for restoring global cooperation in various large-scale distributed systems such as P2P networks [10], for reaching the secure two-party computation in cloud systems [76], or for the influence maximization in various social networking or viral marketing technologies [32].

However, applications of standard tag-based interactions to these technologies without conformity or other potent mechanisms are rather vulnerable to exploitation by cheating free-riders who can fake the cooperative tags to receive the benefits without providing anything in return to their fellow peers. Remaining undetected in distributed, decentralized, and self-organizing systems, such free-riders can pose considerable maintenance costs and cause the ultimate collapse of cooperation in the system. Our present model clearly shows that these earlier limitations can be circumvented by combining tags with the proposed conformity mechanism, enabling the emergence and persistence of global pure altruism in tag-based multi-agent systems. This way, the system does not need to rely on ingroup biased, ethnocentric cooperation, but instead, via tag-mediated interactions and conformity bias, it can generate sustainable levels of global unconditional cooperation. In addition, by employing our discovered pinning control mechanism, systems could be developed that can flexibly and adaptively switch between ingroup favoritism and inter-group cooperation, which may be of particular interest in some technological applications. We hope our present study can serve as the first inspiring step towards this formidable challenge.

## 5. Conclusions

In the present paper, we addressed the question of whether and under what conditions tag-mediated interactions in a networked multiagent system with conformity biased transmission of strategies can generate dominant unconditional cooperation, i.e. pure altruism that can outcompete other strategies. Combining the methods of game theory, evolutionary computing, and agent-based simulation, we instituted a hybrid computational model to investigate the emergence of prosocial behavior attributed to the co-evolutionary [64] interplay between individual strategies and conformity bias. In a series of extensive Monte Carlo simulation experiments, we systematically examined four different variants of our computational spatial model (including the model version with conformity applied to all strategies, with nonconforming altruists, with a random pinning control mechanism, and with long-range connections), which we then additionally compared to the mathematical aspatial model and numerical simulations for the well-mixed population. With this model, which is to the best of our knowledge the first one combining tag-mediated cooperation with conformity, we obtained several non-trivial findings.

Tags can knowingly reduce the complexity of information processing and serve as a reliable trust-inducing mechanism in anonymous one-shot interactions [23], whereas conformity enhances coordinated risk-minimizing behaviors within a group, ensuring that individual payoffs do not drop significantly below the average payoff level [36]. However, none of the two mechanisms alone can push the system into the state of pure unconditional cooperation. Corroborated by recent theoretical [43] and empirical studies [42], our findings suggest that indeed, conformity may not be as strong catalyst of prosociality as has been previously assumed, but that instead, it requires a support from other mechanisms to fully unwind its cooperation-enforcing potency.

We have shown that one such mechanism is the small-world network structure, which is typically known for its fast and reliable signal propagation between network nodes [65,71]. Specifically, our computational experiments and mathematical analyses have revealed that under the majority rule, the nonconformity of co-operators and the small-world spatial structure in combination are the two necessary prerequisites for the stable dominance of altruists in tag-based multiagent systems.

Crucially, our simulations also provide the first evidence that it is possible to significantly amplify the degree of altruism in a tag-based multiagent system by means of a relatively simple control mechanism that needs to be applied only to a small subset of altruistic agents. Evidently, such an ability to enhance pure altruism while simultaneously suppressing all other competing strategies is of paramount importance for any decentralized large-scale system with self-interested units that can easily provoke the fragility of cooperative equilibria.

### Acknowledgments

### Appendix A. Computational experiments: Additional results

Here, we show the results of additional simulations that were conducted to further explore the parameter space. Fig. A.14 shows the strategy frequency $\rho$ as a function of the pinning degree $f_p$ for three different cases of the rewiring probability $p$, confirming thus that irrespective of the underlying rewiring probability of the network, only small fraction $f_p$ of pinning altruists can push the system into the phase of nearly full-altruism.

Fig. A.15 depicts the model variant in which all altruistic agents are initially free from conformity and where cooperative interactions and the preference biases occur via both short and long-range network connections. In Fig. A.16, we show the frequency of strategists as a function of $\delta$ under different network rewiring probability conditions (when interactions among agents remain only local).

In Fig. A.17, we show the evolution of the system under the pinning control mechanism with different proportions of the pinning free altruists. Fig. A.18 shows the typical 2D snapshots of the evolution of strategies under different degrees of the pinning altruists. The dynamical behaviors of various difference measures under varying degrees of the pinning control mechanism are shown in Fig. A.19. Furthermore, in Figs. A.20–A.22, we give the simulation outcomes for the case with nonconformist altruists but with interactions occurring along both local and global network connections, and in Fig. A.23, we show the model version where all agents are conformity biased and interactions occur also along long-range connections.

### Appendix B. Mathematical analysis

*Conformist altruists*

As typically observed in well-mixed populations, we also find that egoism becomes the dominant strategy in the absence of any additional cooperation-promoting mechanisms. According to the rules of the model, there will be $N_s = n_s \times n_c \times n_t = 4 \times 2 \times 4 = 32$ different possible states for each individual in the case that altruists are conformity biased (also see Section 3.1), regardless of whether individuals are living or not, where $n_s$, $n_c$, and $n_t$ denote the number of possible strategies, the number of possible cooperative preferences, and the number of possible tag colors in our model. Thus, unlike in computational experiments, where interactions occur only among living individuals, such restriction is no longer present in the mathematical analysis. Furthermore, we list the payoff matrices in Tables B.1–B.4 which contain $N_s = 32$ payoff elements for each state.

In our tables, we only show the payoff elements for the four types of strategists with the tag color $X = 0$; the positions of the red payoff elements will then shift among the blue elements for the other values of $X$ (i.e. for other tag colors), as described in the caption of Table B.1. More specifically, when a noncooperative altruist interacts with a cooperative ethnocentric individual, we will have the payoff array $X = 1$: $\{0, b, 0, 0\}$ for the tag color $X = 1$. For further tag colors, we have $X = 2$: $\{0, 0, b, 0\}$ and $X = 3$: $\{0, 0, 0, b\}$. Similarly, when a noncooperative altruist interacts with one cooperative cosmopolitan agent $((C, Y, X), X = 0, 1, 2, 3)$, the array will be $X = 1$: $\{b, 0, b, b\}$ for $(A, N, 1)$, $X = 2$: $\{b, b, 0, b\}$ for $(A, N, 2)$, and $X = 3$: $\{b, b, b, 0\}$ for $(A, N, 3)$. The payoff elements in Table B.1 (and subsequent tables) that can be shifted are represented in red color, as elements that are moving among the blue payoff elements. However, when an altruist interacts with an egoist, we see that the payoff array remains $\{0, 0, 0, 0\}$ regardless of the value of $X$, so the color of the payoff elements in Table B.1 for this type of interaction is always black.

Our present analysis is inspired by the work of Traulsen and Schuster [77], which introduced a minimal analytic approach to tag-mediated cooperation. In general, the evolutionary dynamics of cooperation can be analytically described by a set of equations called the replicator dynamics [78,79]. We can thus calculate the mean payoffs from the payoff matrix as [77]

$$\Pi_{(S,P,T)}^t = \sum_{S',P',T'}^{n_s,n_c,n_t} p_{(S',P',T')}^t E_{(S',P',T')}^{(S,P,T)}, \tag{B.1}$$

$$\langle \Pi \rangle^t = \sum_{S,P,T}^{n_s,n_c,n_t} p_{(S,P,T)}^t \Pi_{(S,P,T)}^t \tag{B.2}$$

where $p_{(S',P',T')}^t$ stands for the fraction of individuals in a state $(S', P', T')$ ($S'$, $P'$ and $T'$ correspond to different strategies, cooperation preferences, and tag colors) at time $t$; $E_{(S',P',T')}^{(S,P,T)}$ denotes the payoff elements of $(S, P, T)$ when facing the state $(S', P', T')$. For further details on $E_{(S',P',T')}^{(S,P,T)}$, please see the description in the caption of Table B.1. Moreover, in Eq. (B.2), $\Pi_{(S,P,T)}^t$ is the payoff of the state $(S, P, T)$ with frequency $p_{(S,P,T)}^t$ while $\langle \Pi \rangle^t$ is the mean payoff from the payoff matrix.

Based on Eq. (B.2), and without considering mutation, population structure, the death–birth process (all individuals are instead alive or the number of alive individuals is kept as constant $N$), or conformity, the evolutionary dynamics of the simplest tag-based system could be determined by the following replicator equations [77]:

$$p_{(S,P,T)}^{t+1} = p_{(S,P,T)}^t + h p_{(S,P,T)}^t (\Pi_{(S,P,T)}^t - \langle \Pi \rangle^t), \tag{B.3}$$

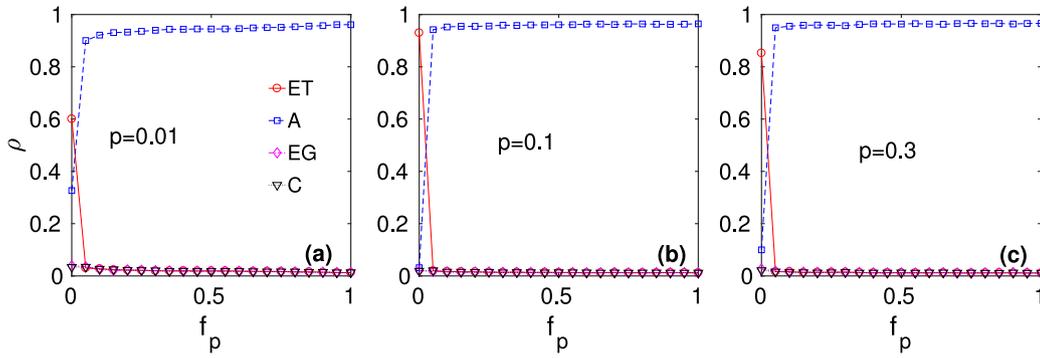**Fig. A.14.** The frequency $\rho$ of the four strategy types as a function of the pinning degree $f_p$ in a 2d WS small-world network with periodic boundary conditions for different rewiring probabilities $p$. The shown results represent the averages taken over $N_r = 50$ independent realizations. The other parameter values were $\delta = 0.6$, $\alpha = 10.0$, $l = 2$, and the rest as described in Methods.
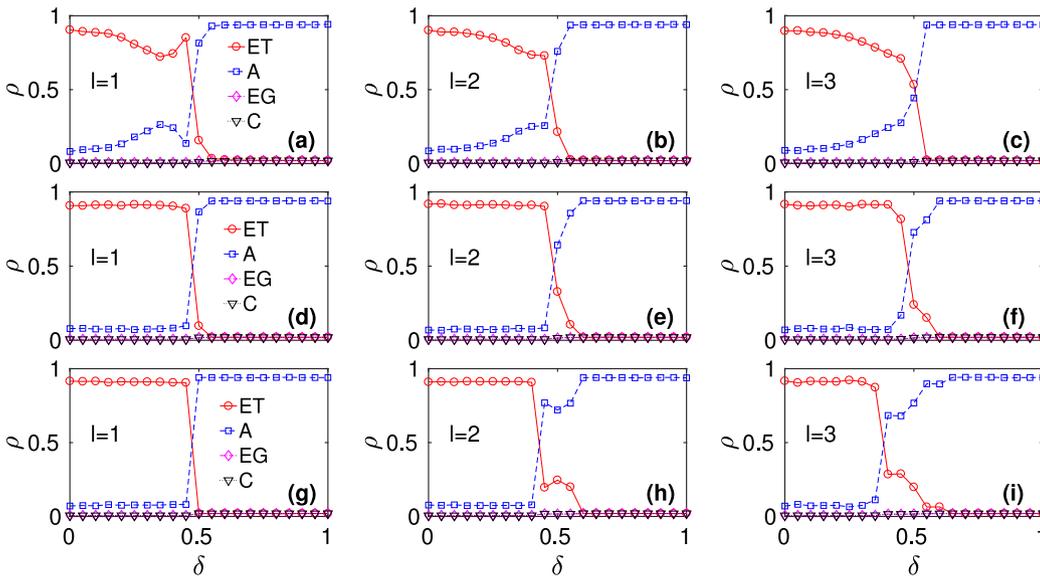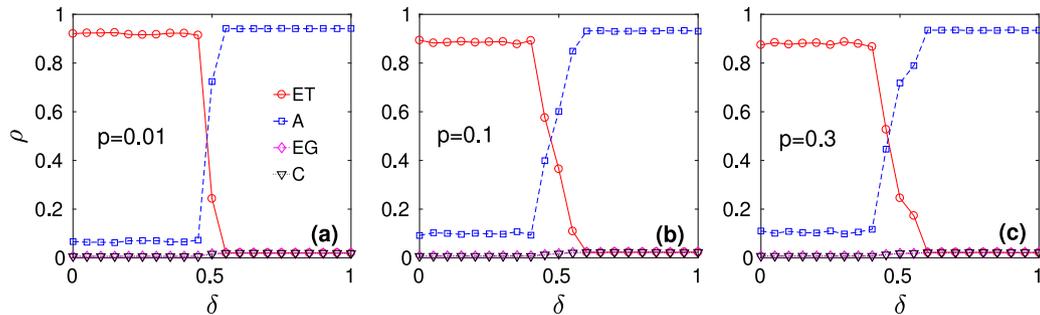


**Fig. A.15.** The frequency $\rho$ of the four strategy types as a function of $\delta$ in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The shown results represent averages taken over $N_r = 50$ independent simulation runs. The other parameter values were $\alpha = 5.0$ (top row of panels), $\alpha = 10.0$ (middle row panels), and $\alpha = 20.0$ (bottom row panels), and the remaining parameters were selected as described in Methods.



**Fig. A.16.** The frequency $\rho$ of the four strategy types as a function of the strength of group influence $\delta$ in a 2d WS small-world network with periodic boundary conditions for the three different network rewiring probability conditions: $p = 0.01$ (a), $p = 0.1$ (b) and $p = 0.3$ (c). The displayed results represent averages taken over $N_r = 50$ independent realizations. The remaining parameter values were taken as $\alpha = 10.0$, $l = 2$, and for the rest of parameters, we selected the baseline values as listed in the Methods section.

where $h$ determines the time scale. Based on Eq. (B.3), we could identify the final states (i. e., attractors) of the simplest system by means of the relationship $\omega_S = \sum_{P,T} p^\infty_{(S,P,T)}$ ($\omega_S$ is the final frequency of strategy $S$). In order to verify and further extend this basic analytical method, we study the simplest death–birth like model for a well-mixed population. Specifically, in each model

generation, each agent serves as a potential donor (gives help of value $b$, but incurs cost $c$) for $M = 8$ other agents chosen at random. Hence, a possible recipient who receives the benefit $b$ is chosen on average $M'$ times. After each generation, for $M'$ times, each agent $i$ compares its fitness $f_i^t$ with the fitness of one randomly chosen agent $j$ $f_j^t$; $j$ will die to leave the position to the new offspring
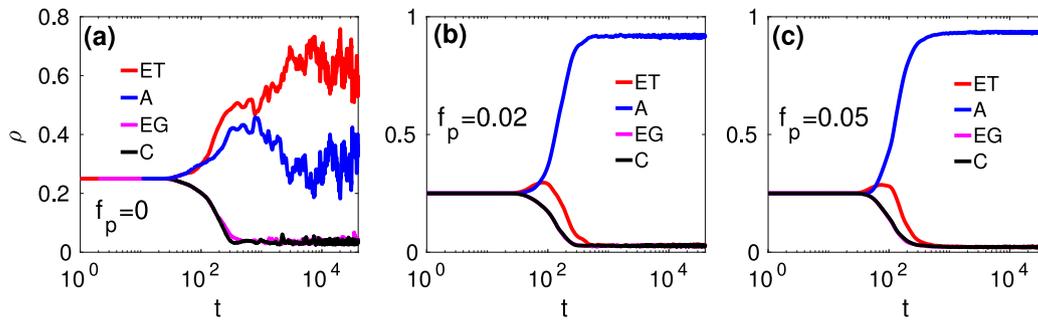
**Fig. A.17.** Evolution of the four strategy types in a 2d WS small-world network with periodic boundary conditions, the rewiring probability $p = 0.01$, and the different proportions of the pinning free altruists. Specifically, in panel (a) $f_p = 0$ i.e., there were no pinning altruists; in (b) $f_p = 0.02$, and $f_p = 0.05$ in panel (c). The remaining parameter values were taken as $\delta = 0.6$ (majority rule), $\alpha = 10.0$, $l = 2$, and the rest as described in the Methods section.
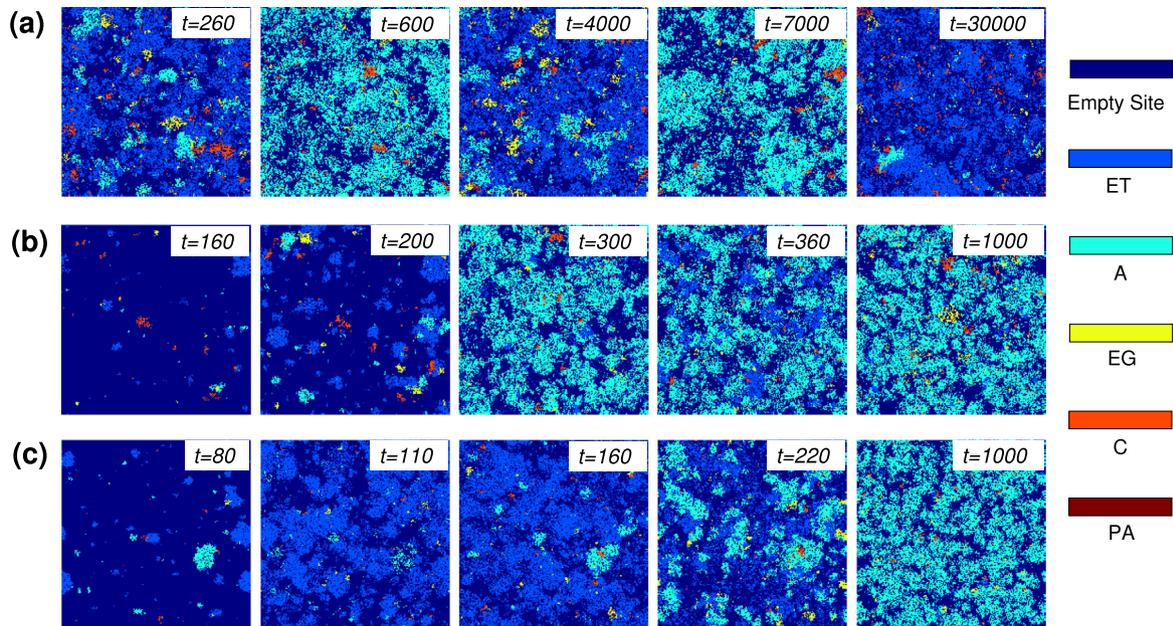


**Fig. A.18.** Typical 2D color snapshots of the evolution of four strategies in a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$, with different degrees of the pinning free altruists. Three different fractions $f_p$ of free altruists were initially set as pinning altruists. Specifically, in panel (a) $f_p = 0$, in (b) $f_p = 0.01$, and in (c) $f_p = 0.05$. The other parameter values were $\delta = 0.6$, $\alpha = 10.0$, $l = 2$, and the rest as detailed in the Methods section.
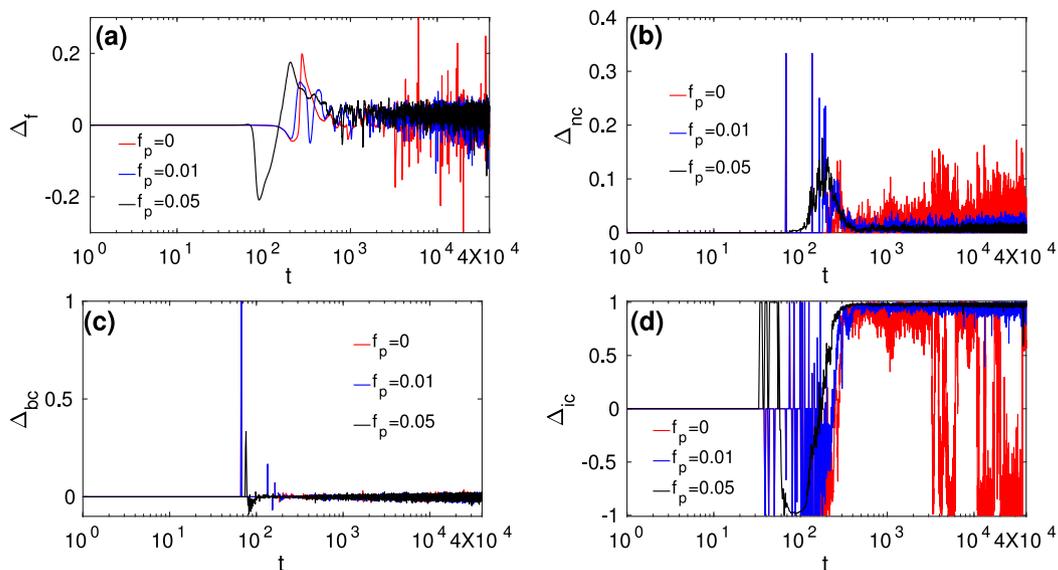


**Fig. A.19.** Dynamical behaviors of $\Delta_f(t)$ (a), $\Delta_{nc}(t)$ (b), $\Delta_{bc}(t)$ (c), and $\Delta_{ic}(t)$ (d) for three different degrees of the pinning control $f_p$. The system evolution was simulated on a 2d WS small-world network with periodic boundary conditions and the rewiring probability $p = 0.01$. The remaining parameter values were $\delta = 0.6$, $\alpha = 10.0$, $l = 2$, and as described in Methods for the rest of the model parameters.
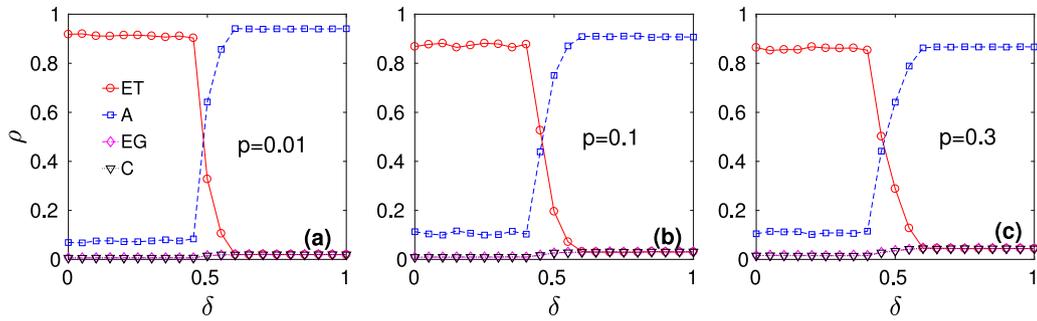
**Fig. A.20.** The frequency $\rho$ of the four strategy types as a function of $\delta$ in a 2d WS small-world network with periodic boundary conditions and the three different rewiring probabilities (a): $p = 0.01$, (b): $p = 0.1$ and (c): $p = 0.3$. The results were obtained by averaging over $N_r = 50$ independent realizations. The other parameter values were $\alpha = 10.0$, $l = 2$, and the rest as detailed in the Methods.
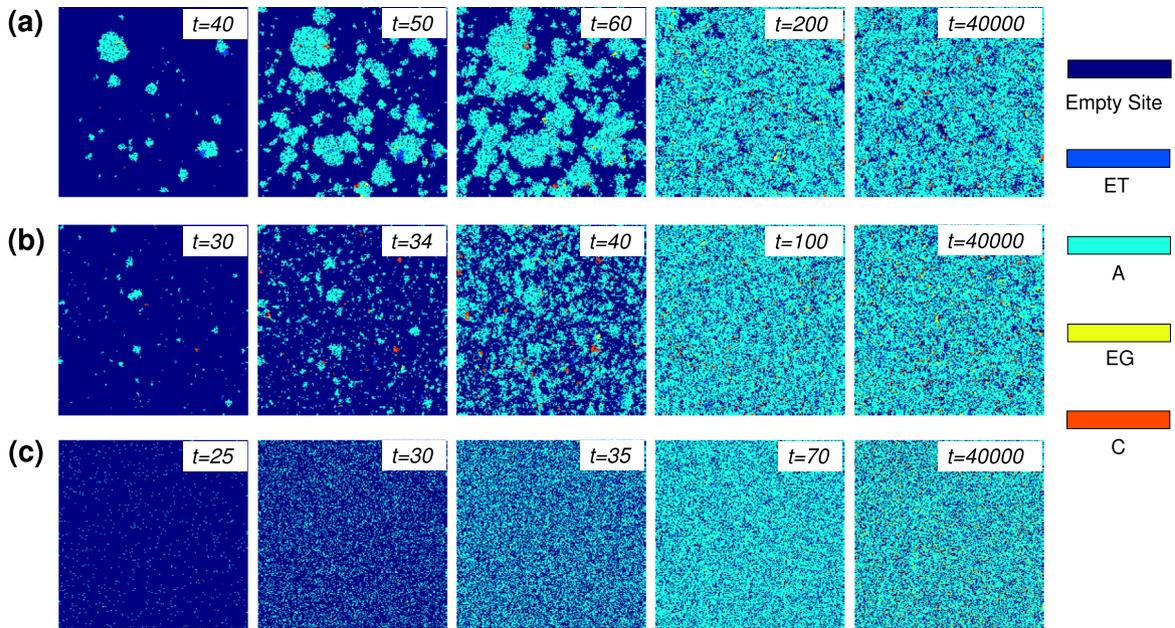


**Fig. A.21.** Typical 2D color snapshots of the four strategies evolving on a 2d WS small-world network with periodic boundary conditions and the three different rewiring probabilities: (a) $p = 0.01$, (b) $p = 0.1$ and (c) $p = 0.3$. In this model version, altruists are free from conformity bias. The other parameter values were $\delta = 0.55$, $\alpha = 10.0$, $l = 2$, and the rest as described in Methods.
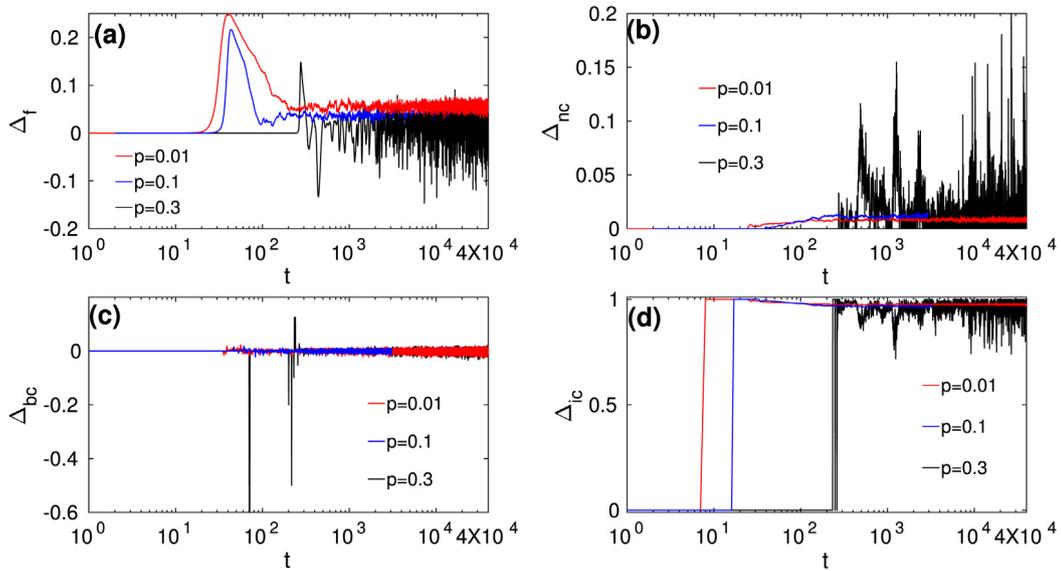


**Fig. A.22.** (a) Dynamical behaviors of $\Delta_f(t)$, (b) $\Delta_{nc}(t)$, (c) $\Delta_{bc}(t)$, and (d) $\Delta_{ic}(t)$ for three different values of $p$. The system evolution was simulated on a 2d WS small-world network with periodic boundary conditions; the conformity bias did not affect altruists in this model version. The other parameter values were $\delta = 0.55$, $\alpha = 10.0$, $l = 2$, and as described in Methods for the rest of model parameters.
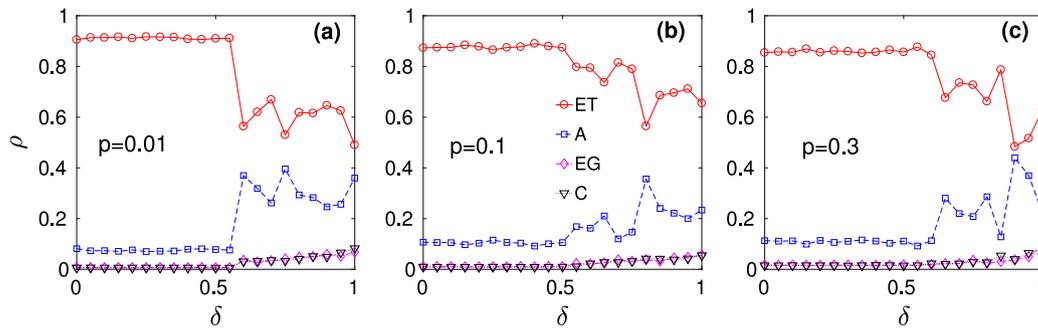
**Fig. A.23.** The frequency $\rho$ of four strategy types as a function of $\delta$ in a 2d WS small-world network with periodic boundary conditions and the three different rewiring probabilities: (a) $p = 0.01$, (b) $p = 0.1$, and (c) $p = 0.3$. In this model version, all agents are conformity biased and interactions occur via both local and global network connections. The results were obtained by averaging over $N_r = 50$ independent realizations. The other parameter values were $\alpha = 10.0$, $l = 2$, and as described in Methods for the remaining model parameters.

**Table B.1**

The payoff elements for noncooperative and cooperative altruists. We note that only the payoff elements for altruists with tag color $X = 0$ are listed; thus, the positions of the red payoff elements would shift among the blue elements with the value of X, i. e. the tag color. For example, for noncooperative altruists we obtain the array of the payoff elements with the tag color $X = 0$: {b, 0, 0, 0} facing cooperative ethnocentric individuals ((Et, Y, X), X = 0,1,2,3). If the tag color is $X = 1$, the array will be {0, b, 0, 0}. Additionally, $E_{(S',P',T')}^{(S,P,T)}$ denotes the payoff elements when an individual in a state $(S, P, T)$ is facing another individual with a state $(S', P', T')$. For example $E_{(ET,Y,2)}^{(A,N,0)} = 0$.

| (A, N, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |
| (A, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
| | -c | -c | -c | -c | b-c | b-c | b-c | b-c |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | -c | -c | -c | -c | b-c | -c | -c | -c |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | -c | -c | -c | -c | -c | -c | -c | -c |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | -c | -c | -c | -c | -c | b-c | b-c | b-c |

*X=0,1,2,3

**Table B.2**

The payoff elements for noncooperative and cooperative ethnocentric agents. We note that only the payoff elements for ethnocentrics with tag color $X = 0$ are listed; the positions of the red payoff elements would then shift among the blue elements with the value of X, i. e. the tag color. For more details, please see the explanation in the caption of Table B.1.

| (ET, N, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |
| (ET, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
| | -c | 0 | 0 | 0 | b-c | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | -c | 0 | 0 | 0 | b-c | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | -c | 0 | 0 | 0 | -c | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | -c | 0 | 0 | 0 | -c | b | b | b |

*X=0,1,2,3

that inherits the fitness, cooperation preference, strategy, and tag color of the parent $i$ if $f_i^t > f_j^t$. For simplicity, the mutation in this death–birth (DB) process is absent. As expected, we found an agreement between the analytical results and the numerical simulations, thereby verifying the analytical method.

Next, we introduce a mutation mechanism into the DB process of the simplest model, where each newborn offspring is subject to mutation. With probability $\mu = 0.005$, a newborn receives both a strategy and a tag color randomly drawn from the set of strategies and tag colors. The Eq. (B.3) can then be developed into

**Table B.3**

The payoff elements for noncooperative and cooperative egoists. According to the rules of model, the payoff elements keep unchanged with the change of cooperative preference of egoists (from being noncooperative to being cooperative). We note that we only list the payoff elements for egoists with tag color $X = 0$; the positions of the red payoff elements would shift among the blue elements with the value of X, i. e. the tag color. For more details, please see the explanation in caption of Table B.1.

| (EG, N, X)/(EG, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |

*X=0,1,2,3

**Table B.4**

The payoff elements for noncooperative and cooperative cosmopolitans. We only list the payoff elements for cosmopolitan agents with the tag color $X = 0$; the positions of the red payoff elements would shift among the blue elements with the value of X, i. e. the tag color. For more details, please see the explanation in caption of Table B.1.

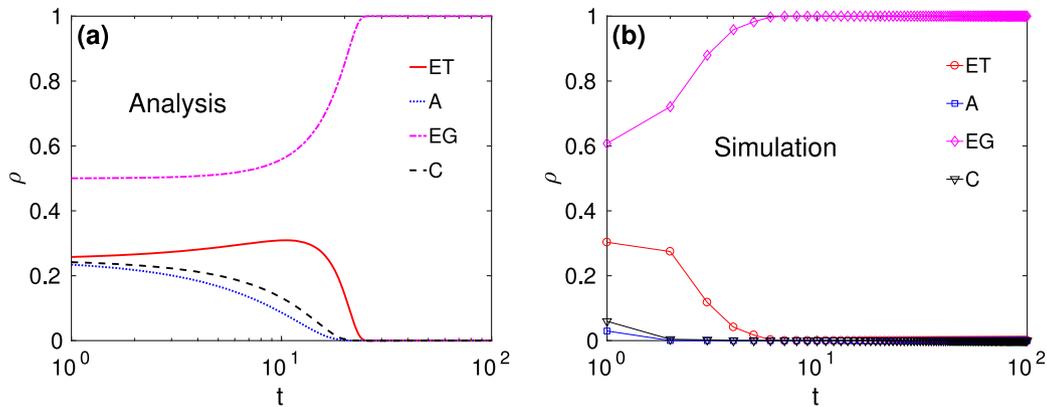| (C, N, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | 0 | 0 | 0 | 0 | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |
| (C, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
| | 0 | -c | -c | -c | b | b-c | b-c | b-c |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | -c | -c | -c | b | -c | -c | -c |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | -c | -c | -c | 0 | -c | -c | -c |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | -c | -c | -c | 0 | b-c | b-c | b-c |

*X=0,1,2,3



**Fig. B.24.** The evolution of the four strategy types in the simplest aspatial cooperation system. The results are obtained from (a) analytical predictions based on Eq. (B.3) and (b) numerical simulations conducted for a well-mixed population ignoring the population structure. The simulated results are obtained by averaging over $N_r = 50$ independent realizations. The size of the system was $N = 4 \times 10^4$, and the other parameters were taken as $\alpha = 10.0$ and $M = 8$.

the following set of equations:

$$p_{(S,P,T)}^{t+1} = p_{(S,P,T)}^t + \Delta_{(S,P,T)_1}^t + \Delta_{(S,P,T)_2}^t + \Delta_{(S,P,T)_3}^t, \qquad \text{(B.4)}$$

$$\Delta_{(S,P,T)_1}^t = h p_{(S,P,T)}^t (\Pi_{(S,P,T)}^t - \langle\Pi\rangle^t)(1 - \mu)$$

$$\Delta_{(S,P,T)_2}^t = -h p_{(S,P,T)}^t (\Pi_{(S,P,T)}^t - \langle\Pi\rangle^t)\mu \frac{15}{16}$$

$$\Delta_{(S,P,T)_3}^t = \sum_{S',T'} h p_{(S',P,T')}^t (\Pi_{(S',P,T')}^t - \langle\Pi\rangle^t)\frac{\mu}{16}$$

Overall, Fig. B.25 verifies the agreement between analytical predictions and our numerical simulations. However, in the limit of $M'$, unlike analytical results, our simulations additionally show

that the implemented mutation mechanism can prevent altruists and ethnocentric agents from going totally extinct. At the same time, we see that the mutation mechanism does not play a decisive role in defining the direction of the system evolution.

Next, we study a well-mixed population with DB dynamics and we consider the interaction rule due to which cooperative acts can be exchanged only among living agents in the population. It is necessary to stress here that it is difficult to provide analytical results to curve the evolution of this system, since the total number of living individuals changes with time. However, our simulations reveal how this interaction rule affects the underlying evolutionary dynamics.
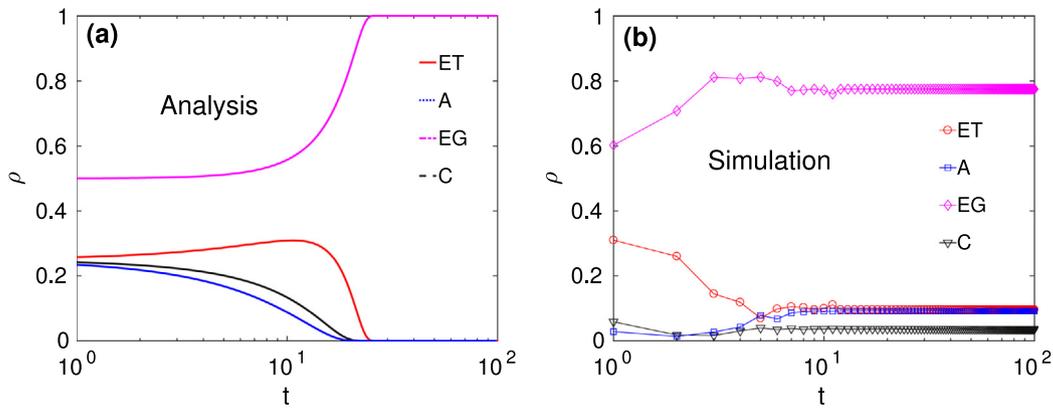
**Fig. B.25.** The evolution of the four types of strategies in the simplest aspatial tag-based system with the mutation mechanism. The shown results are obtained from (a) analytical predictions based on Eq. (B.3) and (b) numerical simulations conducted for a well-mixed population ignoring the spatial effects of the population structure. The simulated results are obtained by averaging over $N_r = 50$ independent realizations. The size of the system was $N = 4 \times 10^4$, and the other parameters were taken as $\alpha = 10.0$ and $M = 8$.
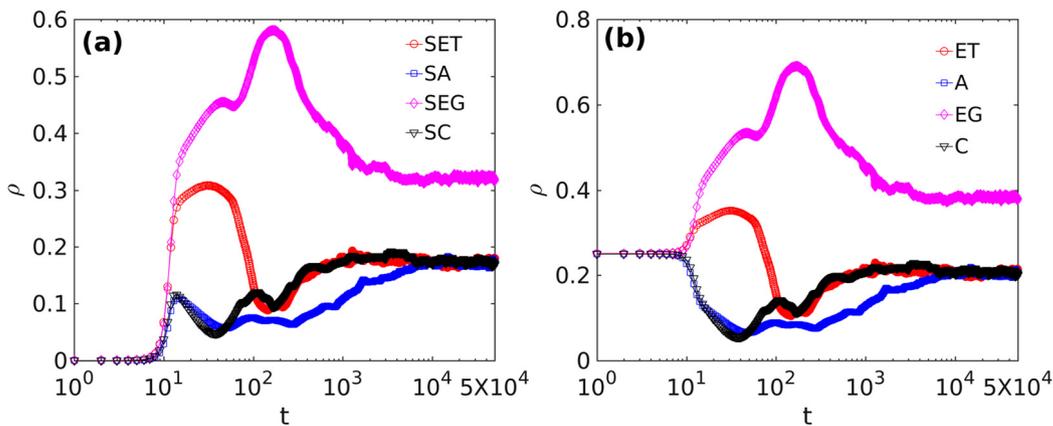


**Fig. B.26.** (a) The evolution of the frequency of living individuals with four different strategies. (b) The evolution of the frequency of total individuals with four different strategies. The size of the simulated system was $N = 4 \times 10^4$, and the other parameter values were taken as $\alpha = 10.0$ and $M = 8$.
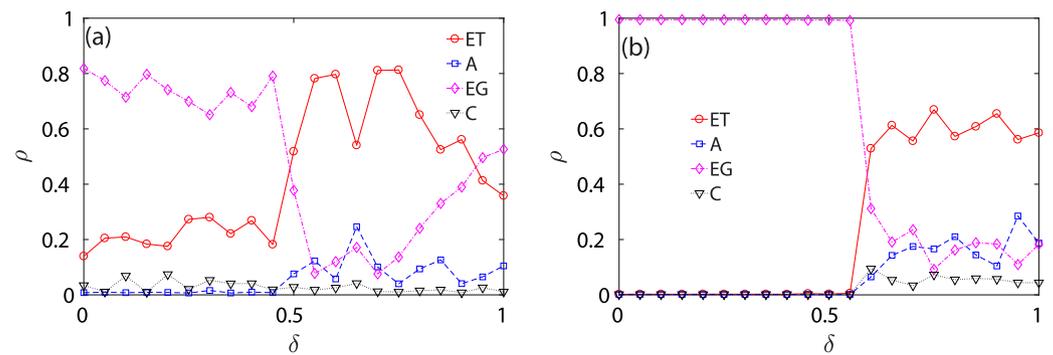


**Fig. B.27.** The frequency $\rho$ of the four strategies as a function of $\delta$ in the aspatial tag-based model with mutation and conformity. Specifically, the interaction rule (i.e. interaction occurs only among living individuals) is absent in (a), but present in (b). The results represent the averages taken over $N_r = 50$ independent realizations for each value of $\delta$. The size of the system was $N = 4 \times 10^4$, and the other parameter values were taken as $\alpha = 10.0$ and $M = 8$.

It can be observed in Fig. B.26 that the egoists are further suppressed due to this interaction rule. Figs. B.25 and B.26 show that the mutation mechanism and the implemented interaction rule both contribute to the decay of egoists which, however, still dominate the population in this aspatial model. Compared to the simulation outcomes shown in Fig. 3(a) of the Results section of this paper, we can conclude that spatial structure indeed plays a decisive role in facilitating the prevalence of ethnocentric agents by enabling them to assort and cluster, and thereby to resist the invasion of egoists (which is also further supported by Fig. 4(a)(b) in the Results section).

Admittedly, we have found that the equations will not be closed again if we introduce the conformity function [37] $\Phi(\omega_Y^t) = \frac{1}{1+\exp(-\alpha(\omega_Y^t-\delta))}$ (where $\omega_Y^t = \sum_{S,T}^{n_s,n_t} p_{(S,Y,T)}^t$ is the total frequency of individuals with the cooperative preference 'Yes') into the equations because of the existence of nonlinear elements. As a result, $\omega_A^t + \omega_{ET}^t + \omega_{EG}^t + \omega_C^t \neq 1.0$. We therefore conclude that it is not possible to provide an analytical treatment of the system with conformity. However, we instead implement numerical
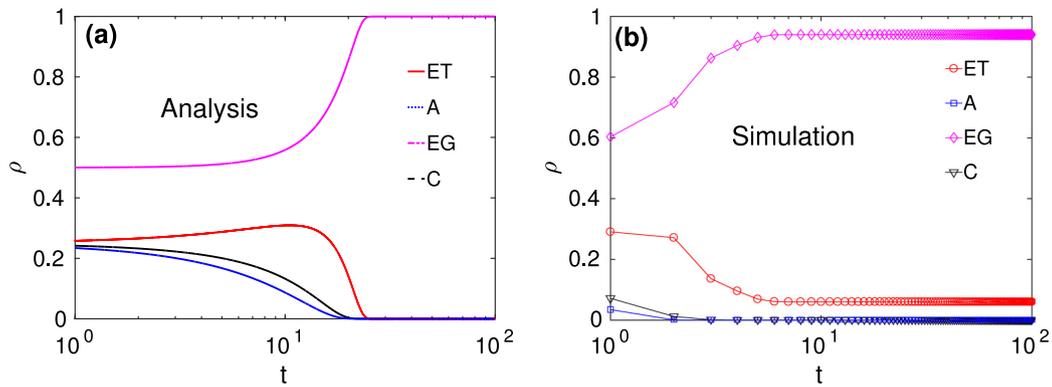
**Fig. B.28.** The evolution of the four strategies in the simplest system with free altruists. Specifically, the results shown are obtained from (a) analytical predictions based on Eq. (B.3) and (b) numerical simulations for the well-mixed condition ignoring the spatial structure of the system. The results correspond to averages taken over $N_r = 50$ independent realizations. The size of the system was $N = 4 \times 10^4$, and the other parameter values were $\alpha = 10.0$ and $M = 8$.
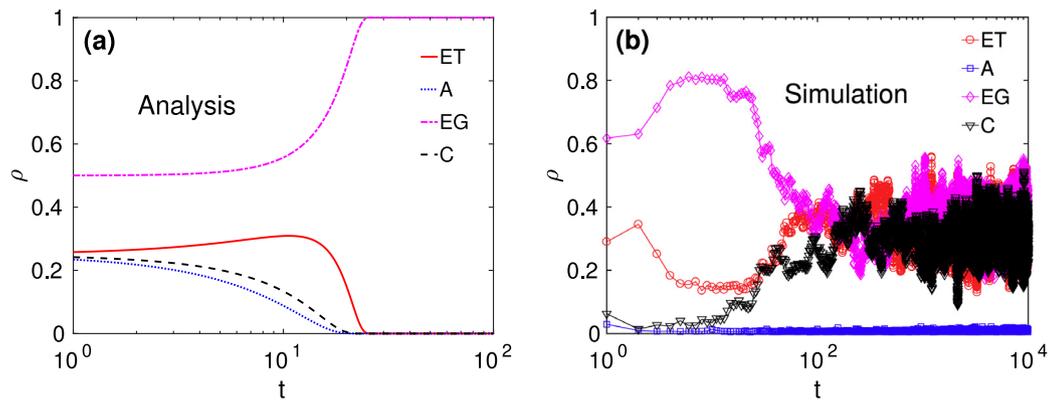


**Fig. B.29.** The evolution of the four strategies in the simplest system with mutation and free altruists. The shown results are obtained from (a) analytical predictions based on Eq. (B.3) and (b) numerical simulations for the well-mixed condition. The results represent the averages taken over $N_r = 50$ independent realizations. The size of the system was $N = 4 \times 10^4$ and the other parameter values were taken as $\alpha = 10.0$ and $M = 8$.



**Fig. B.30.** (a) The evolution of the frequency of living individuals with the four different strategies. (b) The evolution of the frequency of total individuals with the four different strategies. The results represent the averages taken over $N_r = 50$ independent realizations. The size of the system was $N = 4 \times 10^4$, and the other parameter values were taken as $\alpha = 10.0$ and $M = 8$.
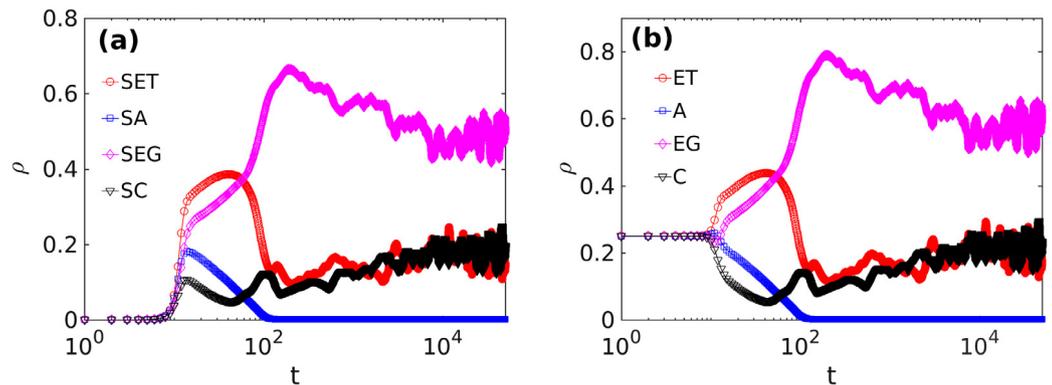
simulations under the well-mixed condition without considering the spatial topology of the system.

Fig. B.27 shows the evolution of the aspatial tag-based cooperation system with conformity. The figure shows that the majority rule ($\delta > 0.5$) can effectively suppress egoists in the population, especially for the case with the implemented interaction rule Fig. B.27(b). Inevitably, altruists remain strongly suppressed and in the minority without the help of the spatial structure. However, compared with Fig. 2 of the Results section, we can

see that when conformity bias is applied to all types of strategists, none of the two key ingredients, i.e. neither conformity bias nor the spatial structure, can tune the evolution in favor of altruists.

*Nonconformist altruists*

Next we shift our attention to the case with nonconformist altruists, which means that altruists' preferences are always 'Yes'
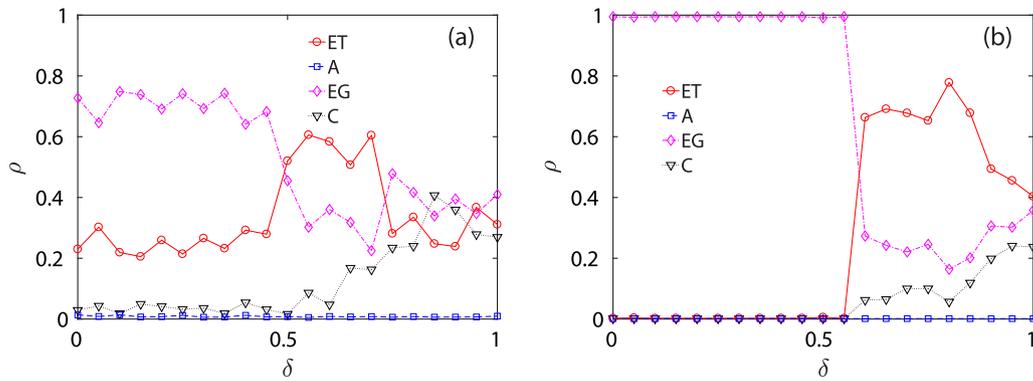
**Fig. B.31.** The frequency of the four strategies as a function of $\delta$. In this model version, all altruists are nonconformist, but conformity of other strategies and mutation are present, and the interaction rule is absent in (a), but present in (b). The results are obtained by averaging over $N_r = 50$ independent realizations for each value of $\delta$. The size of the system was $N = 4 \times 10^4$, $\alpha = 10.0$, and $M = 8$.

**Table B.5**
The payoff elements for noncooperative or cooperative altruists. We note that both noncooperative and cooperative altruists own the same payoff elements since they are free altruists. As before, we also only list the payoff elements for altruists with the tag color $X = 0$; which means that the positions of the red payoff elements would shift among the blue elements with the other value of X, i. e. the tag color.

| (A, N, X)/(A, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | b-c | b-c | b-c | b-c | b-c | b-c | b-c | b-c |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | -c | -c | -c | -c | b-c | -c | -c | -c |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | -c | -c | -c | -c | -c | -c | -c | -c |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | -c | -c | -c | -c | -c | b-c | b-c | b-c |

*X=0,1,2,3

**Table B.6**
The payoff elements for noncooperative and cooperative ethnocentric agents. We only list the payoff elements for ethnocentrics with the tag color $X = 0$; the positions of the red payoff elements would shift among the blue elements with other values of X, i. e. the tag color. For more details, please consult the explanation in the caption of Table B.1.

| (ET, N, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | b | b | b | b | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |
| (ET, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
| | b-c | b | b | b | b-c | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | -c | 0 | 0 | 0 | b-c | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | -c | 0 | 0 | 0 | -c | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | -c | 0 | 0 | 0 | -c | b | b | b |

*X=0,1,2,3

**Table B.7**
The payoff elements for noncooperative and cooperative egoists. According to the rules of our model, the payoff elements remain unchanged with the modification of cooperative preferences of egoists. As before, we only list the payoff elements for egoists with the tag color $X = 0$; the positions of the red payoff elements would shift among the blue elements with other values of X, i. e. with another tag color. For more details, please see the explanation in the caption of Table B.1.

| (EG, N, X)/(EG, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | b | b | b | b | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |

*X=0,1,2,3

**Table B.8**
The payoff elements for noncooperative or cooperative cosmopolitans. We only show the payoff elements for cosmopolitan agents with the tag color $X = 0$; the positions of the red payoff elements would shift among the blue elements with other values of X, i. e. with another tag color. For more details, please consult the explanation in the caption of Table B.1.

| (C, N, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
|---|---|---|---|---|---|---|---|---|
| | b | b | b | b | b | b | b | b |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | 0 | 0 | 0 | b | 0 | 0 | 0 |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | 0 | 0 | 0 | 0 | b | b | b |
| (C, Y, X) | (A ,N, 0) | (A, N, 1) | (A, N, 2) | (A, N, 3) | (A, Y, 0) | (A, Y, 1) | (A, Y, 2) | (A, Y, 3) |
| | b | b-c | b-c | b-c | b | b-c | b-c | b-c |
| | (ET, N, 0) | (ET, N, 1) | (ET, N, 2) | (ET, N, 3) | (ET, Y, 0) | (ET, Y, 1) | (ET, Y, 2) | (ET, Y, 3) |
| | 0 | -c | -c | -c | b | -c | -c | -c |
| | (EG, N, 0) | (EG, N, 1) | (EG, N, 2) | (EG, N, 3) | (EG, Y, 0) | (EG, Y, 1) | (EG, Y, 2) | (EG, Y, 3) |
| | 0 | -c | -c | -c | 0 | -c | -c | -c |
| | (C, N, 0) | (C, N, 1) | (C, N, 2) | (C, N, 3) | (C, Y, 0) | (C, Y, 1) | (C, Y, 2) | (C, Y, 3) |
| | 0 | -c | -c | -c | 0 | b-c | b-c | b-c |

*X=0,1,2,3

and they consequentially always cooperate in the model (for comparison, please see again Section 3.2).

Tables B.5–B.8 show that egoists are still the dominant group in the system. Employing similar methods as in the previous subsection, we firstly proceed with the analytical investigation and numerical simulations for the simplest case without considering mutation, the interaction rule, and the conformity bias. Fig. B.28 illustrates the results from both the analysis and simulations. In comparison to what we presented in Fig. B.24, we observe in Fig. B.28 that with nonconformist free altruists in the model, a certain number of individuals still hold the ethnocentric strategy, in spite of the strong dominance of egoists. This is due to the fact that, besides egoists, the ethnocentric individuals benefit most from the free altruists (see Table B.6). Moreover, the finite number of interactions saves the ethnocentric individuals from overexploitation by egoists.

Fig. B.29 gives the results for the case considering the mutation. As expected, our analytical predictions show that egoists are still evolutionary successful. By contrast, it is surprising to observe in Fig. B.29(b) that mutation begins to play an important role in facilitating the prevalence of other strategies (ethnocentric and cosmopolitan) when the number of interactions is limited in the system. Moreover, we see that due to mutation, the three strategies (ethnocentric, egoistic, and cosmopolitan) start to appear about equally often (though with strong oscillations) in evolutionary time, while altruists remain strongly suppressed.

Furthermore, in Fig. B.30 we show the results for the case in which we additionally consider the earlier mentioned interaction rule. Interestingly, egoists still form the majority of the population in the presence of both mutation and the interaction rule (when altruists are free in the model). Specifically, compared to Fig. B.29, we find that the interaction rule instead plays a negative role (as it actually contributes to the prevalence of egoists).

This interaction rule actually inhibits the exploitation by egoists, however, it also restricts the interactions between living altruists, ethnocentrics, or cosmopolitans. Therefore, both ethnocentric and cosmopolitan agents cannot obtain sufficient help from free altruists, ultimately leading to their decay. We can see in Fig. 9(a) of the Results section that the interaction rule actually contributes to the prevalence of ethnocentric agents, since ethnocentrics and altruists can get together and form mutualistic, coexisting clusters (see e.g. Fig. 10 in the Results section).

In Fig. B.26 and especially in Fig. B.30, we also see that free altruists are actually in a strongly disadvantaged position. Nevertheless, we know from Figs. 6–9 of the Results section that this situation does not hold in the presence of the implemented spatial structure and the conformity bias, especially under the majority rule (see Figs. 6–9 of Results). This finding suggests that spatial structure and conformity are the two indispensable ingredients for the stable dominance of altruists.

Compared to Fig. 6 of the Results section, Fig. B.31 also confirms that both the spatial structure and conformity bias must work at the same time as a condition for the dominance of altruists in a tag-based cooperation model.

## References

[1] M.A. Nowak, R. Highfield, SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed, Free Press, New York, NY, 2011.

[2] H. Yahyaoui, A trust-based game theoretical model for web services collaboration, Knowl.-Based Syst. 27 (2012) 162–169.

[3] Y. Zhu, J. Zhang, Q. Sun, Z.Q. Chen, Evolutionary dynamics of strategies for threshold snowdrift games on complex networks, Knowl.-Based Syst. 130 (2017) 51–61.

[4] M. Mejia, N. Peña, J.L. Muñoz, O. Esparza, M.A. Alzate, A game theoretic trust model for on-line distributed evolution of cooperation in MANETs, J Netw. Comput. Appl. 34 (1) (2011) 39–51.

[5] E. Shakshuki, H. Ghenniwa, M. Kamel, An architecture for cooperative information systems, Knowl.-Based Syst. 16 (2003) 17–27.

[6] X.F. Zha, A knowledge intensive multi-agent framework for cooperative/collaborative design modeling and decision support of assemblies, Knowl.-Based Syst. 15 (2002) 493–506.

[7] N. Dimitriou, A. Polydoros, A. Barnawi, Cooperative schemes for path establishment in mobile ad-hoc networks under shadow-fading, Ad Hoc Netw. 11 (8) (2013) 2556–2566.

[8] T.Y. Wu, W.T. Lee, N. Guizani, T.M. Wang, Incentive mechanism for P2P file sharing based on social network and game theory, J. Netw. Comput. Appl. 41 (2014) 47–55.

[9] F. Liu, X. Li, Y. Ding, H. Zhao, X. Liu, Y. Ma, B. Tang, A social network-based trust-aware propagation model for P2P systems, Knowl.-Based Syst. 41 (2013) 8–15.

[10] D. Hales, B. Edmonds, Applying a socially inspired technique (tags) to improve cooperation in P2P networks, IEEE Trans. Syst. Man Cybernet. 35 (3) (2005) 385–395.

[11] T. Hadzibeganovic, C.Y. Xia, Cooperation and strategy coexistence in a tag-based multi-agent system with contingent mobility, Knowl.-Based Syst. 112 (2016) 1–13.

[12] R.A. Hammond, R. Axelrod, The evolution of ethnocentrism, J. Confl. Resolut. 50 (2006) 926–936.

[13] V.A.A. Jansen, M. van Baalen, Altruism through beard chromodynamics, Nature 440 (2006) 663–666.

[14] N. Masuda, H. Ohtsuki, Tag-based indirect reciprocity by incomplete social information, Proc. Roy. Soc. B 274 (2007) 689–695.

[15] A. Traulsen, M.A. Nowak, Chromodynamics of cooperation in finite populations, PLoS ONE 3 (2007) e270.

[16] T. Antal, H. Ohtsuki, J. Wakeley, P.D. Taylor, M.A. Nowak, Evolution of cooperation by phenotypic similarity, Proc. Natl. Acad. Sci. USA 106 (2009) 8597–8600.

[17] T. Hadzibeganovic, F.W.S. Lima, D. Stauffer, Benefits of memory for the evolution of tag-based cooperation in structured populations, Behav. Ecol. Sociobiol. 68 (2014) 1059–1072.

[18] R.A. Laird, Green beard effect predicts the evolution of traitorousness in the two-tag Prisoner's Dilemma, J. Theoret. Biol. 288 (2011) 84–91.

[19] T. Hadzibeganovic, D. Stauffer, X.P. Han, Randomness in the evolution of cooperation, Behav. Process. 113 (2015) 86–93.

[20] H. Zhang, H. Ye, Role of perception cost in tag-mediated cooperation, Appl. Math. Comput. 279 (2016) 76–89.

[21] C. Efferson, R. Lalive, E. Fehr, The coevolution of cultural groups and ingroup favoritism, Science 321 (2008) 1844–1849.

[22] P. Ramazi, M. Cao, F.J. Weissing, Evolutionary dynamics of homophily and heterophily, Sci. Rep. 6 (2016) 22766.

[23] T. Hadzibeganovic, D. Stauffer, X.P. Han, Interplay between cooperation-enhancing mechanisms in evolutionary games with tag-mediated interactions, Physica A 496 (2018) 676–690.

[24] R.L. Riolo, M.D. Cohen, R. Axelrod, Evolution of cooperation without reciprocity, Nature 414 (2001) 441–443.

[25] M. Chen, L. Wang, S. Sun, J. Wang, C.Y. Xia, Evolution of cooperation in the spatial public goods game with adaptive reputation assortment, Phys. Lett. A 380 (2016) 40–47.

[26] D.G. Rand, M.A. Nowak, Human cooperation, Trends Cogn. Sci. 17 (2013) 413–425.

[27] F.L. Pinheiro, M.D. Santos, F.C. Santos, J.M. Pacheco, Origin of peer influence in social networks, Phys. Rev. Lett. 112 (2014) 098702.

[28] H.X. Yang, Z.-X. Wu, Z. Rong, Y.C. Lai, Peer pressure: Enhancement of cooperation through mutual punishment, Phys. Rev. E 91 (2015) 022121.

[29] A. Whalen, K. Laland, Conformity biased transmission in social networks, J. Theoret. Biol. 380 (2015) 542–549.

[30] S.E. Asch, Studies of independence and conformity: I. A minority of one against a unanimous majority, Psychol. Monographs: Gen. Appl. 70 (1956) 1–70.

[31] J. Henrich, R. Boyd, The evolution of conformist transmission and the emergence of between-group differences, Evol. Human Behav. 19 (1998) 215–241.

[32] F. Morone, H.A. Makse, Influence maximization in complex networks through optimal percolation, Nature 524 (2015) 65–68.

[33] J. Jung, A. Bramson, W.D. Crano, An agent-based model of indirect minority influence on social change and diversity, Soc. Influ. 13 (2018) 18–38.

[34] E.J.C. van Leeuwen, D.B.M. Haun, Conformity without majority? The case for demarcating social from majority influences, Anim. Behav. 96 (2014) 187–194.

[35] P. Siedlecki, J. Szwabinski, T. Weron, The interplay between conformity and anticonformity and its polarizing effect on society, J. Artif. Soc. Soc. Simul. 19 (2016) 9.

[36] J.-J. Wu, C. Li, B.-J. Zhang, R. Cressman, Y. Tao, The role of institutional incentives and the exemplar in promoting cooperation, Sci. Rep. 4 (2014) 6421.

[37] P.-B. Cui, Z.-X. Wu, Impact of conformity on the evolution of cooperation in the prisoner's dilemma game, Physica A 392 (2013) 1500–1509.

[38] F. Alpizar, F. Carlsson, O. Johansson-Stenman, Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica, J. Public Econ. 92 (2008) 1047–1060.

[39] J. Henrich, R. Boyd, Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas, J. Theoret. Biol. 208 (2001) 79–89.

[40] E.C. Nook, D.C. Ong, S.A. Morelli, J.P. Mitchell, J. Zaki, Prosocial conformity: Prosocial norms generalize across behavior and empathy, Person. Soc. Psychol. Bull. 42 (2016) 1045–1062.

[41] H.X. Yang, L. Tian, Enhancement of cooperation through conformity-driven reproductive ability, Chaos Solit. Fract. 103 (2017) 159–162.

[42] A. Romano, D. Balliet, Reciprocity outperforms conformity to promote cooperation, Psychol. Sci. 28 (2017) 1490–1502.

[43] J. Van Cleve, Cooperation, conformity, and the coevolutionary problem of trait associations, J. Theoret. Biol. 396 (2016) 13–24.

[44] K. Eriksson, M. Enquist, S. Ghirlanda, Critical points in current theory of conformist social learning, J. Evol. Psychol. 5 (2007) 67–87.

[45] K. Eriksson, J. Coultas, Are people really conformist-biased? an empirical test and a new mathematical model, J. Evol. Psychol. 7 (2009) 5–21.

[46] J. Peña, H. Volken, E. Pestelacci, M. Tomassini, Conformity hinders the evolution of cooperation on scale-free networks, Phys. Rev. E 80 (2009) 016110.

[47] R.W. Cooper, Coordination Games, Cambridge University Press, Cambridge, 1999.

[48] K. Matsuyama, Explaining diversity: Symmetry-breaking in complementarity games, Amer. Econ. Rev. 92 (2002) 241–246.

[49] B. Skyrms, Dynamics of conformist bias, The Monist 88 (2005) 259–269.

[50] H. Bernhard, U. Fischbacher, E. Fehr, Parochial altruism in humans, Nature 442 (2006) 912–915.

[51] N. Masuda, F. Fu, Evolutionary models of in-group favoritism, F1000Prime Rep. 7 (2015) 27.

[52] M. van Veelen, J. García, D.G. Rand, M.A. Nowak, Direct reciprocity in structured populations, Proc. Natl. Acad. Sci. USA 109 (2012) 9929–9934.

[53] J. García, M. van Veelen, A. Traulsen, Evil green beards: Tag recognition can also be used to withhold cooperation in structured populations, J. Theoret. Biol. 360 (2014) 181–186.

[54] S. Kurokawa, Evolutionary stagnation of reciprocators, Anim. Behav. 122 (2016) 217–225.

[55] E. Bonabeau, Agent-based modeling: Methods and techniques for simulating human systems, Proc. Natl. Acad. Sci. USA 99 (2002) 7280–7287.

[56] V. Grimm, E. Revilla, U. Berger, F. Jeltsch, W.M. Mooij, S.F. Railsback, H.H. Thulke, J. Weiner, T. Wiegand, D.L. DeAngelis, Pattern-oriented modeling of agent-based complex systems: Lessons from ecology, Science 310 (5750) (2005) 987–991.

[57] Z.Q. You, X.P. Han, T. Hadzibeganovic, The role of research efficiency in the evolution of scientific productivity and impact: An agent-based model, Phys. Lett. A 380 (2016) 828–836.

[58] P. Hedström, G. Manzo, Recent trends in agent-based computational research: A Brief Introduction, Sociol. Methods Res. 44 (2) (2015) 179–185.

[59] C. Adami, J. Schossau, A. Hintze, Evolutionary game theory using agent-based methods, Phys. Life Rev. 19 (2016) 1–26.

[60] J.R. Kuhn Jr, J.F. Courtney, B. Morris, E.R. Tatara, Agent-based analysis and simulation of the consumer airline market share for Frontier Airlines, Knowl.-Based Syst. 23 (2010) 875–882.

[61] T. Hadzibeganovic, D. Stauffer, C. Schulze, Agent-based computer simulations of language choice dynamics, Ann. N.Y. Acad. Sci. 1167 (2009) 221–229.

[62] Y. Zhang, H. Chen, J. Lu, G. Zhang, Detecting and predicting the topic change of Knowledge-based Systems: A topic-based bibliometric analysis from 1991 to 2016, Knowl.-Based Syst. 133 (2017) 255–268.

[63] C.P. Roca, J.A. Cuesta, A. Sánchez, Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics, Phys. Life Rev. 6 (2009) 208–249.

[64] Z.G. Huang, Z.-X. Wu, X.J. Xu, J.Y. Guan, Y.H. Wang, Coevolutionary dynamics of networks and games under birth-death and birth dynamics, Eur. Phys. J. B 58 (2007) 493–498.

[65] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440–442.

[66] W.B. Du, H. Zhou, Z. Liu, X.B. Cao, The effect of pinning control on evolutionary Prisoner's Dilemma game, Modern Phys. Lett. B 24 (2010) 2581.

[67] J.R. Riehl, M. Cao, Towards optimal control of evolutionary games on networks, IEEE Trans. Automat. Control 62 (2017) 458–462.

[68] T. Hadzibeganovic, D. Stauffer, C. Schulze, Boundary effects in a three-state modified voter model for languages, Physica A 387 (2008) 3242–3252.

[69] D. Li, W. Li, G. Hu, Z. Zheng, Local conformity induced global oscillation, Physica A 388 (2009) 1243–1248.

[70] M.A. Javarone, Social influences in opinion dynamics: The role of conformity, Physica A 414 (2014) 19–30.

[71] P.-B. Cui, Z.-X. Wu, T. Zhou, Small world yields optimal public goods in presence of both altruistic and selfish cooperators, 2017, arXiv:1711.06878.

[72] M. Hewstone, M. Rubin, H. Willis, Intergroup bias, Annu. Rev. Psychol. 53 (2002) 575–604.

[73] P.-B. Cui, Z.-X. Wu, Selfish punishment with avoiding mechanism can alleviate both first-order and second-order social dilemma, J. Theoret. Biol. 361 (2014) 111.

[74] D. Balliet, L.B. Mulder, P.A. Van Lange, Reward, punishment, and cooperation: A meta-analysis, Psychol. Bull. 137 (2011) 594–615.

[75] R.P. Mann, D. Helbing, Optimal incentives for collective intelligence, Proc. Natl. Acad. Sci. USA 114 (2017) 5077–5082.

[76] T. Li, B.B. Gupta, R. Metere, Socially-conforming cooperative computation in cloud networks, J. Parallel Distrib. Comput. 117 (2018) 274–280.

[77] A. Traulsen, H.G. Schuster, Minimal model for tag-based cooperation, Phys. Rev. E 68 (2003) 046129.

[78] J.W. Weibull, Evolutionary Game Theory, MIT Press, 1997.

[79] J. Hofbauer, K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge University Press, Cambridge, 1998.