



## City Research Online

### City, University of London Institutional Repository

---

**Citation:** Meng, Y., Broom, M. & Li, A. (2023). Impact of misinformation in the evolution of collective cooperation on networks. *Journal of the Royal Society. Interface*, 20(206), 20230295. doi: 10.1098/rsif.2023.0295

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

---

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/31736/>

**Link to published version:** <https://doi.org/10.1098/rsif.2023.0295>

**Copyright:** City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

**Reuse:** Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

---

---

---

City Research Online:

<http://openaccess.city.ac.uk/>

[publications@city.ac.uk](mailto:publications@city.ac.uk)

---

# Impact of misinformation in the evolution of collective cooperation on networks

Yao Meng<sup>1</sup>, Mark Broom<sup>2</sup>, Aming Li<sup>1,3</sup>

September 4, 2023

1. Center for Systems and Control, College of Engineering, Peking University, Beijing 100871, China
2. Department of Mathematics, City, University of London, Northampton Square, London EC1V 0HB, UK
3. Center for Multi-Agent Research, Institute for Artificial Intelligence, Peking University, Beijing 100871, China

**Author for correspondence:** Aming Li; email:amingli@pku.edu.cn

## Abstract

Human societies are organised and developed through collective cooperative behaviours. Based on the information in their environment, individuals can form collective cooperation by strategically changing unfavourable surroundings and imitating superior behaviours. However, facing the rampant proliferation and spreading of misinformation, we still lack systematic investigations into the impact of misinformation on the evolution of collective cooperation. Here we study this problem by classical evolutionary game theory. We find that the existence of misinformation generally impedes the emergence of collective cooperation on networks, although the level of cooperation is slightly higher for weak social cooperative dilemma below a proven threshold. We further show that this possible advantage diminishes as social connections become denser, suggesting that the detrimental effect of misinformation further

increases when ‘social viscosity’ is low. Our results uncover the quantitative effect of misinformation on suppressing collective cooperation, and pave the way for designing possible mechanisms to improve collective cooperation.

## 1 Introduction

The proliferation of misinformation on social networks—information that is incorrect or misleading—has become an escalating concern in public debate and academic research in recent years [1, 2, 3, 4, 5]. In addition to studying the diffusion of misinformation and seeking potential interventions and factors that might remedy misinformation among individuals, researchers further reported that misinformation significantly alters the behaviour of individuals’ decision-making [1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. For instance, a randomised controlled trial shows that misinformation around COVID-19 vaccines induced a decline in the intent of 6.2 percentage points in the UK and 6.4 percentage points in the USA among those who stated that they would definitely accept a vaccine [4]. In 2013, a false tweet claiming that Barack Obama was injured in an explosion sent financial markets into a tailspin, wiping out 130 billion in stock value [15].

Despite that experimental research and survey reveal specific cases where misinformation significantly affects individual decision-making, how misinformation alters the evolution of collective cooperative behaviour has received less attention. To uncover the underlying mechanisms of how cooperative behaviour evolves and persists, researchers have turned to the evolutionary game theory—a powerful framework and canonical paradigm for explaining the evolution of collective cooperation [16, 17, 18, 19, 20, 21]. Focusing on the behavioural dynamics in unstructured populations, recent research has reported that different types of social perception bias play a distinct role in cooperation dynamics [22]. However, how misinformation affects the evolution of cooperation in structured populations—in which interactions between individuals are determined by spatial relationships or social networks [23, 24, 25, 26, 27, 28, 29, 30]—remains widely unknown. Traditionally, individuals are willing to change unfavourable surroundings by detecting other possible locations with higher expected payoffs, where the information that individuals rely on to make decisions is assumed to be completely true [31, 32, 33, 34, 35, 36].

In reality, however, due to the proliferation of misinformation, individuals may receive false information about the alternative locations. Here we first investigate the effect of misinformation on the evolution of social cooperation, and uncover how misinformation alters the migration trajectories of individuals and generally impedes the emergence of cooperation under different population densities.

## 2 Modelling framework

Specifically, we employ the traditional lattice with periodic boundary conditions containing  $N$  sites to represent the spatial social connections, where each site is either empty or occupied by a player. The proportion of the number of individuals to all  $N$  sites describes the density of population ( $d$ ). In a typical two-player prisoner’s dilemma game, each individual could choose either cooperation (C) or defection (D). A cooperator receives “rewards”  $R$  from mutual cooperation, while defectors obtain “punishment”  $P$  from mutual defection. A defector attempting to exploit a cooperator obtains  $T$  and leaves  $S$  to its opponent cooperator. Following the canonical practice [30, 33, 37], a single parameter  $T = b$  is employed here to depict the level of social dilemma for collective cooperation (namely  $R = 1, S = 0, P = 0.1$ ), where  $b$  captures the advantages of defectors over cooperators ( $1 < b < 2$ ).

We start by randomly placing the  $Nd$  individuals on the lattice, where each individual adopts cooperation or defection equally. At the beginning of each round of game, the individual at site  $i$  interacts with their neighbours separately and accumulates the payoff  $\pi_i(s_i)$ , where  $s_i$  indicates the strategy of the individual at site  $i$ . Then a randomly selected individual  $i$  on the lattice will decide whether to migrate to other empty sites based on the social information. For different lattices, the migration range of each individual contains  $n_M$  sites, consisting of the current site and the sites whose distance from this site is 1 or larger (figure 1a). The individual evaluates the fictitious expected payoffs  $\tilde{\pi}_j(s_i)$  for any empty site  $j$  in its migration range by hypothetically interacting with neighbours of the site  $j$  using its current strategy  $s_i$ . Then the individual will choose to migrate to  $j$  with the highest expected payoff if  $\tilde{\pi}_j(s_i) > \pi_i(s_i)$ . Afterwards, the individual imitates the strategy of the best performing neighbour (namely, the neighbour having the highest

payoff) if its own current payoff is lower.

Compared to the scenario where individuals obtain true information about the strategies of their neighbours around the alternative locations, misinformation appears within a certain fixed area. In the area with misinformation, any defector is superficially recognised as a cooperator when individuals explore the expected payoffs of the corresponding empty sites (figure 1b). Here, the system state may correspond to the configuration of each site—either empty or occupied by a cooperator or defector, and misinformation leads to a misperception of the local state, which may further affect the individual action—whether and where to migrate. Note that the real strategy of each individual can be correctly recognised when individuals imitate the strategies from their neighbours after migration.

### 3 Results

We first explore how misinformation alters the evolution of social cooperation on lattices. Intriguingly, the environment with misinformation even presents a slight enhancement of the frequency of cooperators at weak dilemma (small  $b$ ) compared to the scenario with completely true information (figure 2a). Since all players in the area with misinformation are recognised as cooperators when individuals evaluate the fictitious expected payoffs of the empty sites, they crowd into the misinformation area to pursue higher payoffs, and thus increase the local population density (namely, the density of individuals) (figure 2b). Besides, we show that the level of cooperation in the misinformation area is also improved (figure 2c).

Previous research uncover that the emergence of collective cooperation relies on the formation of stable cooperative clusters [17, 18, 19, 23, 24, 25]. To investigate the mechanism by which misinformation affects the evolution of collective cooperation, we analyse whether stable cooperative clusters—in which the cooperators neither move nor change their strategy—can be formed during the evolutionary process with misinformation. We define the migration-stable state to represent the state that individuals do not migrate after evaluating the expected payoffs of all empty sites among their migration range. Mathematically, in this state, for any individual at site  $i$ , the expected payoff  $\tilde{\pi}_j(s_i)$  at any

empty site  $j$  within its migration range is no more than its current payoff  $\pi_i(s_i)$ , namely

$$\tilde{\pi}_j(s_i) \leq \pi_i(s_i).$$

An individual can retain its strategy if its own payoff is higher than that of all neighbours (otherwise it should have the same strategy as the neighbour who has the highest payoff), or the neighbour has a higher payoff possessing the same strategy as  $i$ . Specifically, when the system reaches migration-stable state and all individuals satisfy one of the following conditions

$$\begin{cases} s_j \in \{C, D\}, & \text{if } \pi_{j \in \mathcal{N}_i}(s_j) \leq \pi_i(s_i) \\ s_j = s_i, & \text{if } \pi_{j \in \mathcal{N}_i}(s_j) > \pi_i(s_i) \end{cases},$$

where  $\mathcal{N}_i$  indicates the neighbour set of  $i$ . In such case, the evolutionary process ends with a constant frequency of cooperators, and the stable clusters of cooperators are maintained (electronic supplementary material, figures S1 and S2).

Next, we provide the condition for maintaining stable clusters of cooperators for different lattices (figure 2d). During the evolutionary process, to prevent defectors from invading cooperative clusters, the defector (site  $i$  in figure 2d) on the boundary of a cooperative cluster should have a lower payoff than the cooperator (site  $j$  in figure 2d) with full cooperative neighbours, which means

$$n_C^D b + (k - n_C^D) P < kR. \quad (3.1)$$

Here  $n_C^D$  represents the number of cooperative neighbours for the defector ( $i$ ), and  $k$  captures the number of neighbours for each site on the lattice. The left side represents the payoff of a defector interacting with  $n_C^D$  cooperators and  $k - n_C^D$  defectors, and the right side represents the payoff of a cooperator interacting with  $k$  cooperative neighbours. To satisfy equation (3.1), a defector who achieves the maximum payoff can have  $k - 1$  cooperative neighbours, which results in the threshold  $b^* = (kR - P)/(k - 1)$ , below which the defector on the boundary of a stable cooperative cluster can have no more than  $k - 1$  cooperative neighbours (table 1). From the perspective of forming and maintaining stable cooperative clusters, we further uncover that when the ‘‘temptation’’  $b$  is weak (i.e., weak dilemma of collective cooperation), namely lower than the threshold  $b^*$ , the level of cooperation is

maintained in an environment with misinformation and is even slightly higher than that with completely true information (figure 2a).

We next reveal the mechanism by which collective cooperation in the environment with misinformation is maintained under weak social dilemma by analysing the fate of a defector during the evolutionary process. For the defector on the boundary of a stable cooperative cluster (who are allowed to have  $k - 1$  cooperative neighbours when  $b < b^*$ ), the misinformation misleads it to a site with full neighbours in the area with misinformation to pursue higher payoff (see the left panel in figure 2e). However, this brings an even lower payoff for the defector in the subsequent interaction and drives it to continue its migration to a new site (see the middle panel in figure 2e). Finally, the defector imitates the cooperative neighbour who has a higher payoff, and a stable cooperative cluster is thus established (see the right panel in figure 2e). For the whole population, when there are no empty sites with  $k$  neighbours, individuals with  $k - 1$  cooperative neighbours will then stay put, leading to the emergence of the migration-stable state (electronic supplementary material, videos S1–S6). At this point, the number of defectors has decreased since defectors with  $k - 1$  cooperative neighbours have been driven to migrate from original sites, and ultimately imitate their cooperative neighbours in new locations (figure 2e). We numerically confirm the threshold  $b^*$  for a slight promotion of collective cooperation with misinformation at a weak social dilemma on the triangular, square and hexagonal lattices with different population densities (figure 3a). Furthermore, we check the robustness of our results for the Prisoner’s dilemma game with two parameters (electronic supplementary material, figure S3), where a slight promotion of cooperation with misinformation is shown in a region consistent with our proven threshold.

Despite that the level of collective cooperation is slightly enhanced when  $b < b^*$ , we next show that misinformation impedes the emergence of cooperation in strong dilemma of social cooperation ( $b \geq b^*$ ), especially for dense social connections (figure 3a). The unending migration induced by misinformation prevents the formation of stable cooperative clusters. From equation (3.1), we know that when  $b \geq b^*$ , the defectors on the boundary of a stable cooperative cluster should have no more than  $k - 2$  cooperative neighbours in order to sustain the stable cooperative cluster. Due to the existence of the misinformation, individuals keep migrating to sites with  $k - 1$  or  $k$  neighbours, which actually

can not sustain the formation of stable cooperative clusters. Therefore, the migration of individuals never stops due to the considerable amount of empty sites with  $k - 1$  neighbours, leaving no possibility to form stable cooperative clusters (electronic supplementary material, figures S4–S6, videos S7–S9). During the evolutionary game process with misinformation, mobile individuals will fill the entire misinformation area, resulting in the general reduction of the level of cooperation for large  $b$  (figure 3a). As the social connections become denser (from the left to right panel in figure 3a), the threshold  $b^*$  decreases (see equation (3.1)) and then the negative impact of misinformation is enlarged.

Furthermore, we show the impact of population density on the emergence of collective cooperation. The high population density further shrinks the alternative empty sites for migration in the environment with true information. With strong dilemmas, individuals tend to stay put and imitate their neighbours, which accelerates the disintegration of cooperative clusters (electronic supplementary material, figures S7–S9, videos S10–S12). This explains why the frequency of cooperation decreases drastically as population density increases in strong dilemmas (see the black lines in figure 3b, and figure 3c). In contrast, in the environment with misinformation, the variation of population density causes a relatively mild change in the level of collective cooperation since the mobile individuals will fully fill the misinformation environment irrespective of the population density when  $b \geq b^*$  (see the orange lines in figure 3b, and figure 3d).

By numerical simulations and theoretical analysis, we check the robustness of our results for random regular networks (figure 4a, see Appendix). Figure 4b shows the theoretical approximation and simulation results at different population densities on random regular graphs, where we demonstrate that the emergence of social cooperation is generally impeded by misinformation both theoretically and numerically. Compared to previous results for lattices, the level of cooperation in the scenario with misinformation is slightly improved on the random regular networks at high densities and strong dilemmas. This is due to the shorter characteristic path length on random regular graphs compared with lattices, where a small migration range is enough for individuals to migrate out of the area with misinformation to form mobile cooperative clusters in a single step.

In the opposite modelling setting where any cooperator is superficially recognised as a defector in the area with misinformation, we find that misinformation further impedes the

emergence of cooperation (electronic supplementary material, figure S10). In this scenario, misinformation generally impedes the migration of individuals because they estimate lower expected payoffs from sites in the misinformation area. This further hinders the formation of cooperative clusters in this area and leads to a decrease in the frequency of cooperators. Our results uncover that misinformation significantly affects the evolution of cooperation on spatial structures by influencing the migration of individuals.

What is the effect of misinformation on the evolution of cooperation when mutual cooperation is also a Nash equilibrium? To answer this question, we study the evolutionary process with a general Stag Hunt game [38]. We find that stable cooperative clusters are easy to maintain in this scenario since cooperators obtain more payoffs with a cooperative opponent relative to defectors ( $R > T$ ). However, misinformation hinders the formation of cooperative clusters when the advantage of cooperation is weak, resulting in a lower level of cooperation (electronic supplementary material, figure S11). This result further confirms that misinformation is generally detrimental to cooperation in a wide range of games, population densities and social structures.

## 4 Discussion

Overall, misinformation misleads individuals in their decision and leads to ineffective migration, which ultimately inhibits the formation of stable clusters of collective cooperation in a wide range of social dilemma and population densities (electronic supplementary material, figures S12–S17). A promising direction for future research is the effects of misinformation on coevolutionary games where both the environment and strategies evolve over time [30, 39]. Indeed, during the evolution of strategies, there are many factors in the environment that can feed back into the behaviour of individuals [39, 40]. Considering that previous studies have assumed that individuals receive completely true information from the environment, the coevolutionary games affected by misinformation may be a more realistic reflection of how individual behaviour—in particular, collective social cooperation—evolves in practice.

We study the evolution of cooperation where misinformation leads to a false perception of strategy for individuals and directly affects the migration of individuals. Indeed,

whether misinformation impedes or favours the emergence of cooperation generally depends on how one operationalises misinformation. This could be further analysed based on our model. For example, misinformation around COVID-19 vaccines may concern false information about the state of the world. In the environment where individuals prefer living with higher vaccination rates, following our modelling framework, the state of the system can be the configuration of each individual’s strategy—choosing whether or not to vaccinate so that an individual can estimate the local vaccination rate for each available place, and the action of an individual is whether and where to migrate. Therefore, the misinformation about whether the individual vaccinates will influence the estimation of the local vaccination rate of each available place, which further affects the action (migrating or staying put) of the individual.

Nevertheless, to analyse the impact of misinformation on collective behaviours, there are many different frameworks [22, 41]. For instance, a previous study exploring the effect of perception bias on cooperation dynamics where individuals adapt their behaviour through the best response based on their estimates of the current distribution of strategies [22]. Future investigations on the model design beyond the action of individuals in the environment with misinformation are warranted.

Another promising direction for future research lies in differentiating types of misinformation. There are several ways to model how misinformation blurs individuals’ perceptions. On the one hand, perceptions are multiple and it is possible to perceive partial but true information. A previous study has shown that veridical strategies can be driven to extinction by non-veridical strategies during the evolutionary game process [42]. Therefore, it is important to explore how the perception of partially true information affects the evolution of cooperation. On the other hand, instead of blurring the perception of the level of cooperation [22], misinformation may affect the individual perception of the game type [41]. It would be practical to further study the effect of this type of misinformation on cooperation dynamics in structured populations.

Current efforts are mainly devoted to exploring effective interventions to mitigate the negative effects of misinformation on individuals as well as societies, by understanding what the misinformation is, how it operates, and designing the mechanism for suppressing the spread of misinformation [2, 3, 5]. However, the challenge remains to create a new

system of safeguards to put an end to the impact of misinformation in collective social cooperation [2]. Our study sheds new light on the effect of misinformation on the emergence of a high level of collective cooperation. We uncover that defectors on the boundary of the cooperative cluster can be misled to locations with lower real payoffs, providing the opportunity to maintain stable cooperative clusters by controlling the “temptation” of social defection below a proven threshold. Our findings suggest that even in environments with misinformation, effective intervention and organisation by reducing the profit of defectors can be designed to dwindle the potential impact of misinformation in the emergence of collective social cooperation.

**Authors’ Contributions.** Y.M. and A.L. designed the model and performed the research. All authors wrote the paper.

**Competing Interests.** We declare we have no competing interests.

**Data Accessibility.** This article does not contain any additional data.

**Funding.** This work was supported by the National Key Research and Development Program of China under Grant no. 2022YFA1008400; National Natural Science Foundation of China under Grant no. 62173004 and Beijing Nova Program (Z211100002121105); and the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955708.

**Acknowledgements.** We thank the anonymous referees for their constructive comments. We thank the members of the Li Lab for helpful comments and discussions.

## Appendix

### (a) Pair approximation

We represent the probability of a site occupied by cooperator (C), defector (D) or remained vacant (V) as  $p_c$ ,  $p_d$  and  $p_v$  separately, and denote the probability of finding a cooperator with a neighbouring cooperator by  $p_{cc}$ . Similarly, the probabilities of the configuration of CD, CV, DD pairs are defined as  $p_{cd}$ ,  $p_{cv}$  and  $p_{dd}$ , respectively. Together with these four pair configuration probabilities, the evolutionary dynamics driven by migration can be illustrated by the corresponding expected change rate, which is calculated by considering all neighbouring configurations at original sites and migrated sites. The trajectory where a cooperator at site  $i$ , with  $n_c^o$  cooperative neighbours and  $n_d^o$  defective neighbours, stays

put and imitates its defective neighbour occurs with the probability

$$\begin{aligned} P_{C \rightarrow D}^{i \rightarrow i}(n_c^o, n_d^o; n_c^m, n_d^m) &= p_c P_C(n_c^o, n_d^o) P_{M \max}^{(C)}(n_c^m, n_d^m) \\ &\quad \mathbf{1}(\tilde{\pi}_C(n_c^m, n_d^m) \leq \pi_i) p_{\max}^d, \end{aligned}$$

where  $P_C(n_c^o, n_d^o)$  represents the probability of a cooperator having  $n_c^o$  cooperators and  $n_d^o$  defectors out of  $k$  neighbours at its original site.  $\tilde{\pi}_C(n_c^m, n_d^m)$  captures the fictitious expected payoff of a cooperator interacting with  $n_c^m$  cooperators and  $n_d^m$  defectors.  $\mathbf{1}(\cdot)$  is the indicator function.  $P_{M \max}^{(C)}(n_c^m, n_d^m)$  indicates the probability of an empty site with  $n_c^m$  cooperative and  $n_d^m$  defective neighbours having the maximum fictitious payoff for a cooperator in its migration range, which is

$$\begin{aligned} P_{M \max}^{(C)}(n_c^m, n_d^m) &= P_V(n_c^m, n_d^m) \\ &\quad \frac{([\mathbf{P}(\tilde{\pi}_j \leq \tilde{\pi}_C(n_c^m, n_d^m))]^{n_V} - [\mathbf{P}(\tilde{\pi}_j < \tilde{\pi}_C(n_c^m, n_d^m))]^{n_V})}{\sum_{n_c^u, n_d^u} P_V(n_c^u, n_d^u) \mathbf{1}(\tilde{\pi}_C(n_c^u, n_d^u) = \tilde{\pi}_C(n_c^m, n_d^m))}, \end{aligned}$$

where  $n_V$  is the number of empty sites in the migration range. Analogously,  $P_V(n_c^m, n_d^m)$  represents the probability of a vacant site having  $n_c^m$  cooperators and  $n_d^m$  defectors out of  $k$  neighbours. We define  $p_{\max}^d$  as the probability that the highest payoff is from a defector among the new neighbours, which is

$$p_{\max}^d = \mathbf{P}(\pi_l < \pi_h, \pi_i < \pi_h, \forall l \in \mathcal{N}_j, l \neq h, s_h = D).$$

The probability of a cooperator at site  $i$  (with  $n_c^o$  cooperative neighbours and  $n_d^o$  defective neighbours) migrating to an empty site  $j$  (with  $n_c^m$  cooperative neighbours and  $n_d^m$  defective neighbours) and keeping cooperation is

$$\begin{aligned} P_{C \rightarrow V \rightarrow C}^{i \rightarrow j}(n_c^o, n_d^o; n_c^m, n_d^m) &= p_c P_C(n_c^o, n_d^o) \\ &\quad P_{M \max}^{(C)}(n_c^m, n_d^m) \mathbf{1}(\tilde{\pi}_C(n_c^m, n_d^m) > \pi_i) (p_{\max}^o + p_{\max}^c), \end{aligned}$$

where  $p_{\max}^o = \mathbf{P}(\pi_l \leq \pi_i, \forall l \in \mathcal{N}_j)$  denotes the probability that individual has the highest payoff among its new neighbours, and

$$p_{\max}^c = \mathbf{P}(\pi_l < \pi_h, \pi_i < \pi_h, \forall l \in \mathcal{N}_j, l \neq h, s_h = C)$$

indicates the probability that the highest payoff is from a cooperator among the new neighbours. Analogously, the probability of a cooperator at site  $i$  (with  $n_c^o$  cooperative neighbours and  $n_d^o$  defective neighbours) migrating to empty site  $j$  (with  $n_c^m$  cooperative neighbours and  $n_d^m$  defective neighbours) and imitating its new defective neighbour is

$$\begin{aligned} P_{C \rightarrow V \rightarrow D}^{i \rightarrow j}(n_c^o, n_d^o; n_c^m, n_d^m) &= p_c P_C(n_c^o, n_d^o) P_{M \max}^{(C)}(n_c^m, n_d^m) \\ &\quad \mathbf{1}(\tilde{\pi}_C(n_c^m, n_d^m) > \pi_i) p_{\max}^d. \end{aligned}$$

The probability for a defector changing its neighbour configuration is similarly obtained as  $P_{D \rightarrow C}^{i \rightarrow i}$ ,  $P_{D \rightarrow V \rightarrow C}^{i \rightarrow j}$  and  $P_{D \rightarrow V \rightarrow D}^{i \rightarrow j}$ .

## (b) Stable frequency of cooperators

The expected change rate of the pair configuration probability is given by

$$\begin{aligned}
\dot{p}_{cc} &= \frac{2}{Nk} \sum_{n_c^o, n_d^o; n_c^m, n_d^m} (P_{C \rightarrow D}^{i \rightarrow i} \Delta n_{cc}^{C \rightarrow D} + P_{D \rightarrow C}^{i \rightarrow i} \Delta n_{cc}^{D \rightarrow C} \\
&\quad + P_{C \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cc}^{C \rightarrow V \rightarrow D} + P_{C \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cc}^{C \rightarrow V \rightarrow C} \\
&\quad + P_{D \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cc}^{D \rightarrow V \rightarrow D} + P_{D \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cc}^{D \rightarrow V \rightarrow C}), \\
\dot{p}_{cd} &= \frac{1}{Nk} \sum_{n_c^o, n_d^o; n_c^m, n_d^m} (P_{C \rightarrow D}^{i \rightarrow i} \Delta n_{cd}^{C \rightarrow D} + P_{D \rightarrow C}^{i \rightarrow i} \Delta n_{cd}^{D \rightarrow C} \\
&\quad + P_{C \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cd}^{C \rightarrow V \rightarrow D} + P_{C \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cd}^{C \rightarrow V \rightarrow C} \\
&\quad + P_{D \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cd}^{D \rightarrow V \rightarrow D} + P_{D \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cd}^{D \rightarrow V \rightarrow C}), \\
\dot{p}_{cv} &= \frac{1}{Nk} \sum_{n_c^o, n_d^o; n_c^m, n_d^m} (P_{C \rightarrow D}^{i \rightarrow i} \Delta n_{cv}^{C \rightarrow D} + P_{D \rightarrow C}^{i \rightarrow i} \Delta n_{cv}^{D \rightarrow C} \\
&\quad + P_{C \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cv}^{C \rightarrow V \rightarrow D} + P_{C \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cv}^{C \rightarrow V \rightarrow C} \\
&\quad + P_{D \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{cv}^{D \rightarrow V \rightarrow D} + P_{D \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{cv}^{D \rightarrow V \rightarrow C}), \\
\dot{p}_{dd} &= \frac{2}{Nk} \sum_{n_c^o, n_d^o; n_c^m, n_d^m} (P_{C \rightarrow D}^{i \rightarrow i} \Delta n_{dd}^{C \rightarrow D} + P_{D \rightarrow C}^{i \rightarrow i} \Delta n_{dd}^{D \rightarrow C} \\
&\quad + P_{C \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{dd}^{C \rightarrow V \rightarrow D} + P_{C \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{dd}^{C \rightarrow V \rightarrow C} \\
&\quad + P_{D \rightarrow V \rightarrow D}^{i \rightarrow j} \Delta n_{dd}^{D \rightarrow V \rightarrow D} + P_{D \rightarrow V \rightarrow C}^{i \rightarrow j} \Delta n_{dd}^{D \rightarrow V \rightarrow C}),
\end{aligned}$$

where  $\Delta n_{cc}^{C \rightarrow D}$  represents the variation of the number of CC edges caused by a cooperator staying put and imitating the strategy from a defector. Analogously,  $\Delta n_{cc}^{C \rightarrow V \rightarrow C}$  ( $\Delta n_{cc}^{C \rightarrow V \rightarrow D}$ ) corresponds to the variation of the number of CC edges caused by a cooperator migrating to an empty site and being a cooperator (defector). And the variations for the number of CD, CV and DD edges can be similarly obtained. With initial pair probabilities  $p_{cc} = p_c^2, p_{cd} = p_c p_d, p_{cv} = p_c p_v, p_{dd} = p_d^2$ , the stable frequency of cooperators can be solved through numerical iterations. And the frequency of cooperators  $p_c$  at the system equilibrium can be obtained by the summation of equilibrium pair probabilities  $p_{cc}, p_{cd}$  and  $p_{cv}$ .

## (c) Theoretical analysis for evolutionary games with misinformation

The evolutionary process with misinformation takes  $p_{cc}^t, p_{cd}^t, p_{cv}^t, p_{dd}^t, p_{cc}^m, p_{cd}^m, p_{cv}^m, p_{dd}^m, p_{dv}^m$  to describe the dynamical system, which represent the corresponding pair configuration probabilities in true and misinformation areas respectively. Thus,  $P_{D^x \rightarrow C^x}^{i \rightarrow i}, P_{D^x \rightarrow V^x \rightarrow C^x}^{i \rightarrow j}, P_{D^x \rightarrow V^x \rightarrow D^x}^{i \rightarrow j}$  are calculated subsequently, where  $x \in \{t, m\}$ , describing the migration and strategy change in area with true and misinformation. The expected change rate of pair

probabilities of CC, CD, CV and DD pairs are calculated analogously, and the stable values for pair probabilities are obtained through numerical iterations.

## References

- [1] Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D. & Rand, D. G. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021).
- [2] Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J. & Zittrain, J. L. The science of fake news. *Science* **359**, 1094–1096 (2018).
- [3] Brashier, N. M., Pennycook, G., Berinsky, A. J. & Rand, D. G. Timing matters when correcting fake news. *Proc. Natl Acad. Sci. USA* **118**, 2–4 (2021).
- [4] Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K. & Larson, H. J. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nat. Hum. Behav.* **5**, 337–348 (2021).
- [5] Scheufele, D. A. & Krause, N. M. Science audiences, misinformation, and fake news. *Proc. Natl Acad. Sci. USA* **116**, 7662–7669 (2019).
- [6] Pennycook, G. & Rand, D. G. The psychology of fake news. *Trends. Cogn. Sci.* **25**, 388–402 (2021).
- [7] van der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat. Med.* **28**, 460–467 (2022).
- [8] West, J. D. & Bergstrom, C. T. Misinformation in and about science. *Proc. Natl Acad. Sci. USA* **118**, e1912444117 (2021).
- [9] Tandoc, E. C. The facts of fake news: A research review. *Sociol. Compass* **13**, 1–9 (2019).
- [10] Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).
- [11] Holme, P. & Rocha, L. E. Impact of misinformation in temporal network epidemiology. *Netw. Sci.* **7**, 52–69 (2019).

- [12] Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E. & Quattrocioni, W. The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559 (2016).
- [13] Kopp, C., Korb, K. B. & Mills, B. I. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to "fake news". *PLoS ONE* **13**, 1–35 (2018).
- [14] Pennycook, G. & Rand, D. G. Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nat. Commun.* **13**, 1–12 (2022).
- [15] Rapoza, K. Can 'fake news' impact the stock market. *Forbes* <https://www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/> (26 February 2017).
- [16] Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
- [17] Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics* (Cambridge, UK: Cambridge University Press, 1998).
- [18] Nowak, M. A. *Evolutionary Dynamics: Exploring the Equations of Life* (Cambridge, MA: Harvard University Press, 2006).
- [19] Levin, S. A. Collective cooperation: from ecological communities to global governance and back. In *Unsolved Problems in Ecology*, 311–317 (Princeton, NJ: Princeton University Press, 2020).
- [20] Sigmund, K. *The Calculus of Selfishness* (Princeton, NJ: Princeton University Press, 2010).
- [21] Rapoport, A., Chammah, A. M. & Orwant, C. J. *Prisoner's Dilemma: A study in Conflict and Cooperation* (Michigan, MI: Michigan University Press, 1965).
- [22] Santos, F. P., Levin, S. A. & Vasconcelos, V. V. Biased perceptions explain collective action deadlocks and suggest new mechanisms to prompt cooperation. *Isience* **24**, 102375 (2021).
- [23] Santos, F. C. & Pacheco, J. M. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.* **95**, 098104 (2005).
- [24] Hauert, C. & Doebeli, M. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646 (2004).

- [25] Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505 (2006).
- [26] Taylor, P. D., Day, T. & Wild, G. Evolution of cooperation in a finite homogeneous graph. *Nature* **447**, 469–472 (2007).
- [27] Santos, F. C., Santos, M. D. & Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454**, 213–216 (2008).
- [28] Allen, B., Lippner, G., Chen, Y.-T., Fotouhi, B., Momeni, N., Yau, S.-T. & Nowak, M. A. Evolutionary dynamics on any population structure. *Nature* **544**, 227–230 (2017).
- [29] Li, A., Broom, M., Du, J. & Wang, L. Evolutionary dynamics of general group interactions in structured populations. *Phys. Rev. E* **93**, 022407 (2016).
- [30] Li, A., Zhou, L., Su, Q., Cornelius, S. P., Liu, Y.-Y., Wang, L. & Levin, S. A. Evolution of cooperation on temporal networks. *Nat. Commun.* **11**, 1–9 (2020).
- [31] Vainstein, M. H. & Arenzon, J. J. Disordered environments in spatial games. *Phys. Rev. E* **64**, 051905 (2001).
- [32] Vainstein, M. H., T.C. Silva, A. & Arenzon, J. J. Does mobility decrease cooperation? *J. Theor. Biol.* **244**, 722–728 (2007).
- [33] Helbing, D. & Yu, W. The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc. Natl Acad. Sci. USA* **106**, 3680–3685 (2009).
- [34] Sicardi, E. A., Fort, H., Vainstein, M. H. & Arenzon, J. J. Random mobility and spatial structure often enhance cooperation. *J. Theor. Biol.* **256**, 240–246 (2009).
- [35] Chen, X., Szolnoki, A. & Perc, M. Risk-driven migration and the collective-risk social dilemma. *Phys. Rev. E* **86**, 1–8 (2012).
- [36] Meloni, S., Buscarino, A., Fortuna, L., Frasca, M., Gómez-Gardeñes, J., Latora, V. & Moreno, Y. Effects of mobility in a population of prisoner’s dilemma players. *Phys. Rev. E* **79**, 3–6 (2009).
- [37] Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nature* **359**, 826–829 (1992).
- [38] Skyrms, B. *The stag hunt and the evolution of social structure* (Cambridge, UK: Cambridge University Press, 2004).

- [39] Perc, M. & Szolnoki, A. Coevolutionary games—a mini review. *BioSystems* **99**, 109–125 (2010).
- [40] Pacheco, J. M., Traulsen, A. & Nowak, M. A. Coevolution of strategy and structure in complex networks with dynamical linking. *Phys. Rev. Lett.* **97**, 258103 (2006).
- [41] Barfuss, W. & Mann, R. P. Modeling the effects of environmental and perceptual uncertainty using deterministic reinforcement learning dynamics with partial observability. *Phys. Rev. E* **105**, 034409 (2022).
- [42] Mark, J. T., Marion, B. B. & Hoffman, D. D. Natural selection and veridical perceptions. *J. Theor. Biol.* **266**, 504–515 (2010).

Table 1: Condition for maintaining stable cooperative clusters. In order to maintain the stable cooperative clusters, given the maximum number of cooperative neighbours for the defector ( $n_C^D$ ) at the boundary of the cooperative cluster, the range of the “temptation”  $b$  for different lattices is solved according to equation (3.1).

$b$	$n_C^D = 5$	$n_C^D = 4$	$n_C^D = 3$	$n_C^D = 2$	$n_C^D = 1$
Triangular lattice ( $k = 3$ )	–	–	–	[1, 1.45)	[1.45, 2)
Square lattice ( $k = 4$ )	–	–	[1, 1.3)	[1.3, 1.9)	[1.9, 2)
Hexagonal lattice ( $k = 6$ )	[1, 1.18)	[1.18, 1.45)	[1.45, 1.9)	[1.9, 2)	–

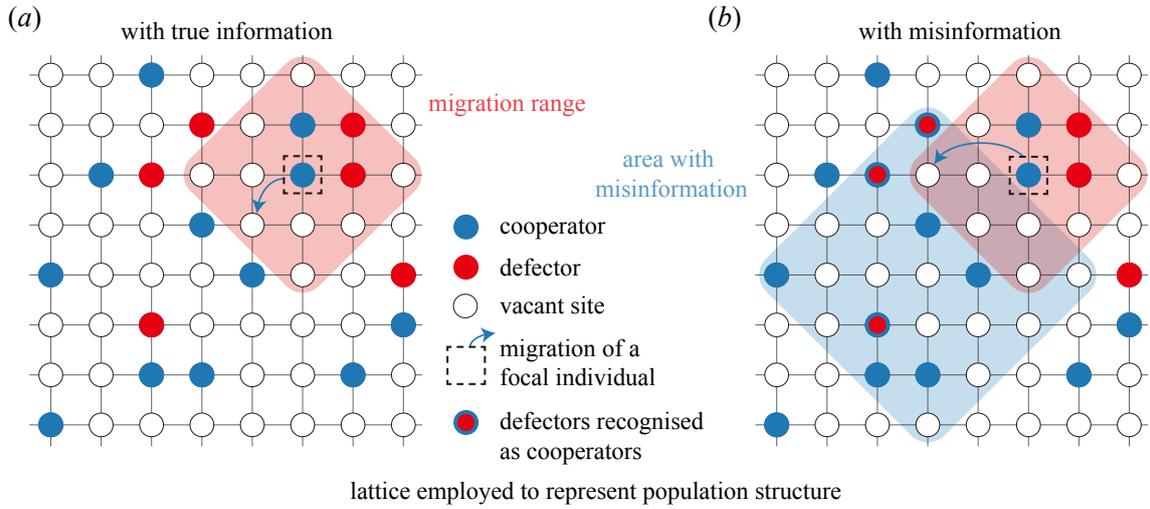


Figure 1: Illustration on the individual migration in areas with true information and misinformation. (a) Individuals are located on the lattice, where each site is either occupied by a cooperator (blue), or a defector (red) or remains vacant (white). In each round of game, individuals interact with their neighbours and accumulate payoffs. Then, individuals are allowed to migrate inside its migration range (red shaded area), which contains  $n_M = 13$  sites. To determine the location for migration, the individual (circled by dashed square) evaluates expected payoff of each empty site by fictitiously interacting with the site's neighbours. Afterwards, the individual chooses to move to the site with two cooperative neighbours (pointed by the solid arrow), which has the highest expected payoff among all available locations (see modelling framework). In the case that several sites have the same highest expected payoff, a random one is selected. And individuals will stay put if the highest expected payoff is lower than the individual's current payoff. (b) The area with misinformation is filled with blue. Defectors in this area are recognised as cooperators when individuals evaluate the fictitious expected payoffs in this area (circles with a red face and blue edge). Note that the site around with a cooperator and two defectors (pointed by the solid arrow) has the highest expected payoff and becomes the best choice in the individual's migration range.

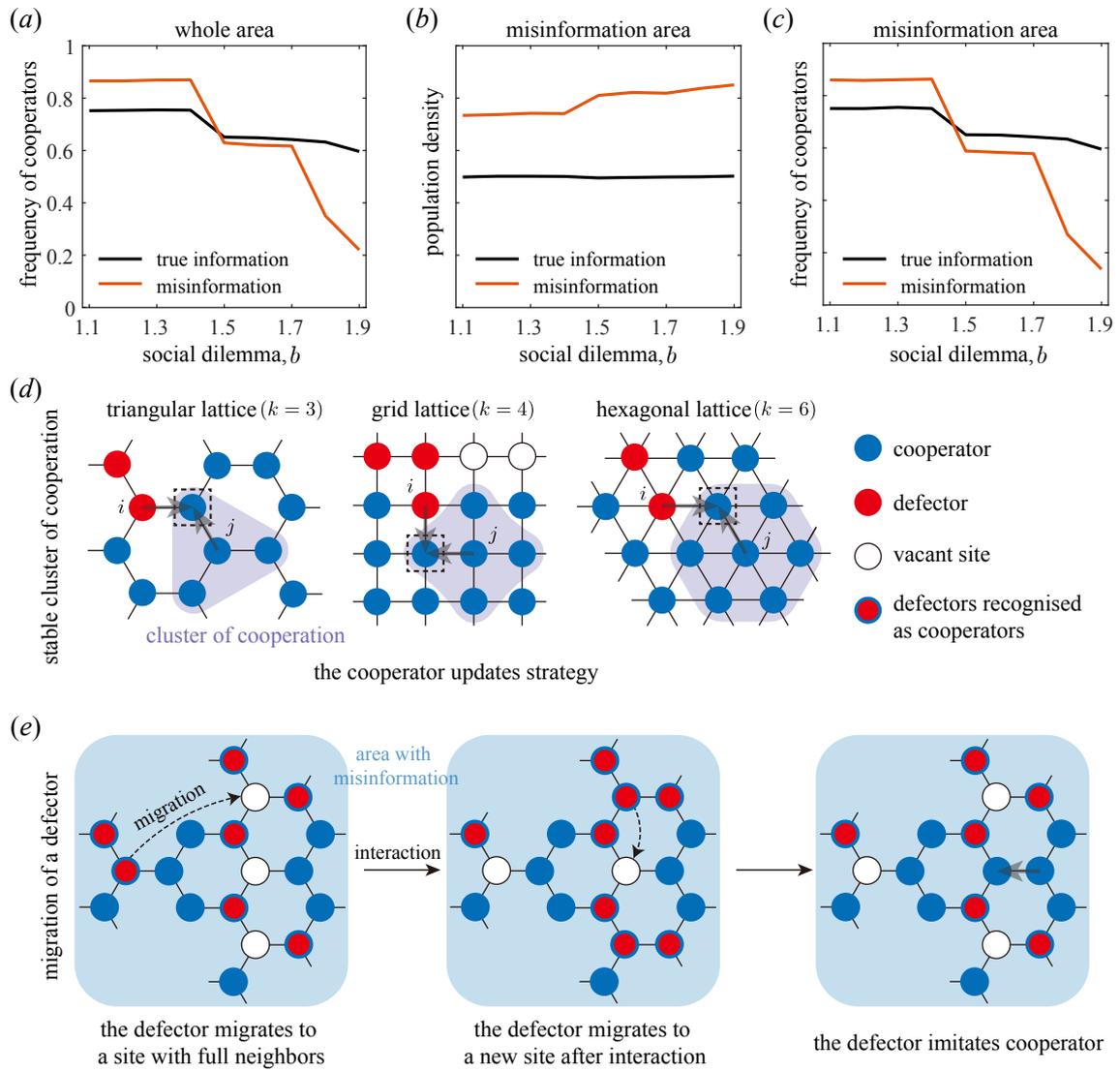


Figure 2:

Figure 2: With misinformation, collective cooperation is promoted at weak social dilemma but inhibited at strong social dilemma. (a) For different levels of social dilemma ( $b$ ), we plot the stable frequency of cooperators on a triangular lattice. The frequency of cooperators with misinformation (orange) exceeds that with true information (black) when  $b$  is small. As individuals are allowed to migrate during evolution, the population density in area with misinformation is higher than that with true information at all range of  $b$  as shown in (b). (c) The frequency of cooperators in the area with misinformation is also higher than that with true information in the same area at weak social dilemma, which is consistent with (a). The stable frequency of cooperators in (a)–(c) are obtained over 60 independent runs, where the average frequency in each run is obtained over 2000 rounds after a transient time of  $2 \times 10^5$  rounds. In the environment with misinformation, the configurations for forming stable cooperative clusters at weak dilemma for triangular, square and hexagonal lattices are illustrated in (d), where a defector could have  $k - 1$  cooperative neighbours with  $k$  being the maximum number of neighbours. Taking the left panel as an example, the defector with two cooperative neighbours and the cooperator with full cooperative neighbours compete (indicated by gray arrow) for the site in the dashed box. To retain the cooperative cluster (purple shaded area), the payoff of the defector should be lower than the cooperator with full cooperative neighbours (see equation (3.1)). (e) Illustration of the process for a defector in the misinformation area to migrate and become a cooperator. The defector with two cooperative neighbours is misled to migrate to the site with full defective neighbours, which results in a reduction of the payoff after a round of game and further migration as shown in the middle panel. After moving to a site with a cooperative neighbour with a higher payoff, the defector imitates the cooperator in the right panel.

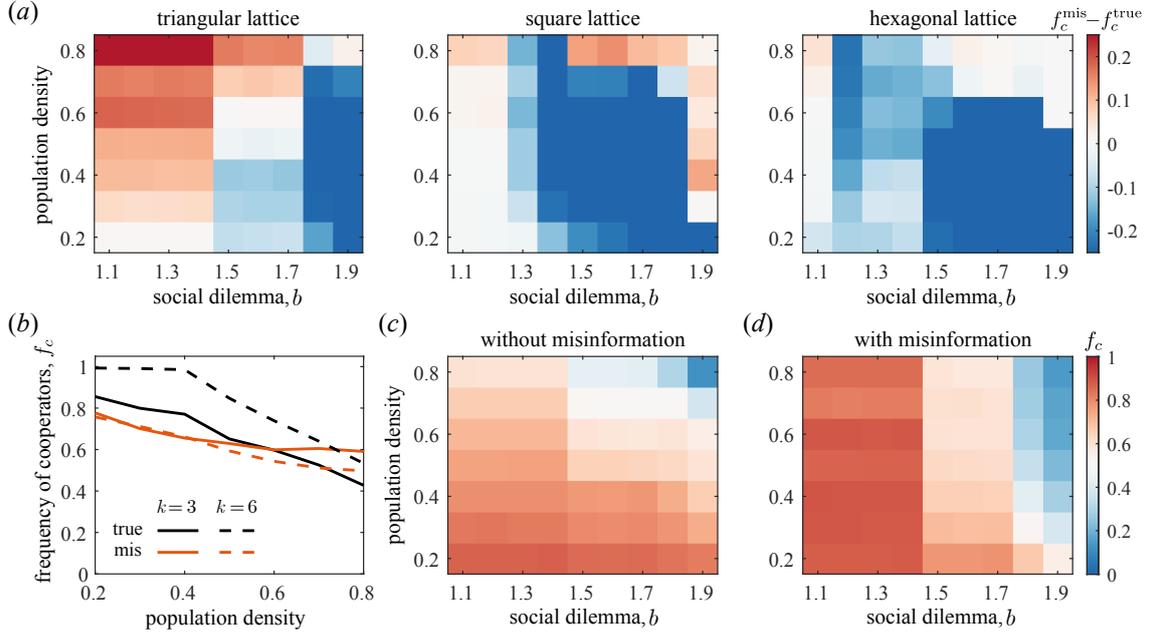


Figure 3: Effect of population density on collective cooperation with misinformation and true information. We compare the frequency of cooperators during the evolution with misinformation ( $f_c^{\text{mis}}$ ) and that with true information ( $f_c^{\text{true}}$ ) on triangular ( $k = 3$ ), square ( $k = 4$ ), hexagonal ( $k = 6$ ) lattices with periodic boundaries with population density ranging from 0.2 to 0.8. The difference with individual migration range of  $n_M = 121$  is plotted in (a), where the red colour indicates  $f_c^{\text{mis}} > f_c^{\text{true}}$  and blue indicates  $f_c^{\text{mis}} < f_c^{\text{true}}$ . (b) We show the stable frequency of cooperators ( $f_c$ ) through population density with true information (black) and misinformation (orange) on triangular (solid line) and hexagonal lattices (dashed line) at  $b = 1.5$ . Furthermore, we plot  $f_c$  for the evolutionary process on a triangular lattice with true information in (c) and that with misinformation in (d) over different population densities and levels of social dilemma, which indicates that the variations of population density generally leave a weak effect on the evolution of cooperation with misinformation. The total number of nodes on triangular, square and hexagonal lattices are 2380, 2401, 2312 respectively.

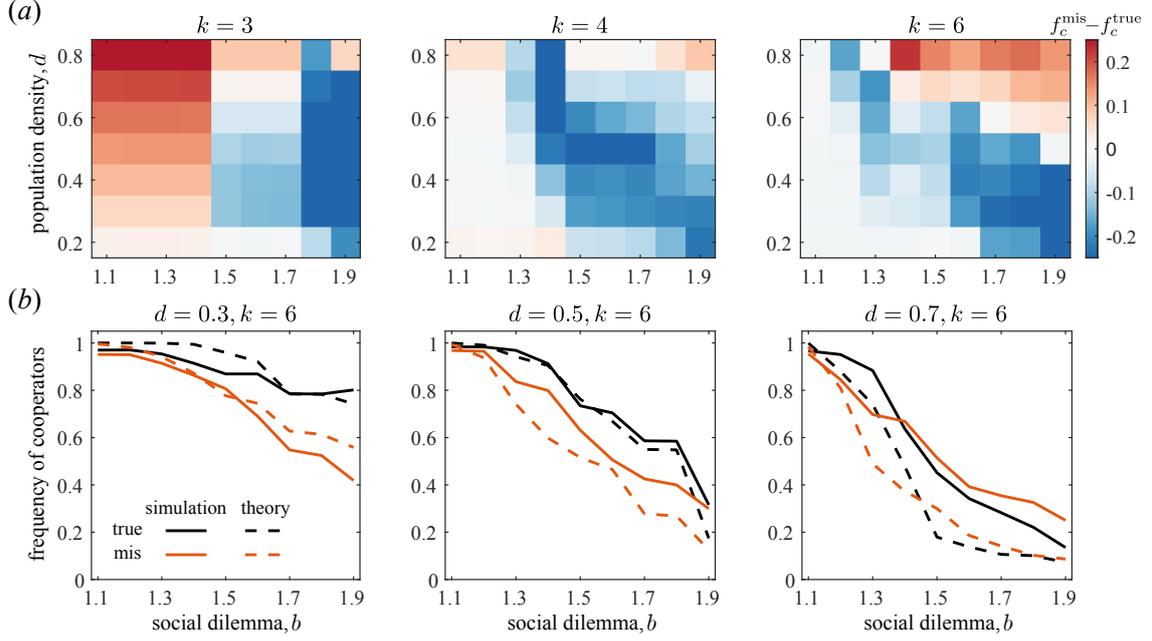


Figure 4: Our results hold true for large random regular social connections. (a) We show the heatmap of the difference between frequency of cooperators with misinformation ( $f_c^{\text{mis}}$ ) and true information ( $f_c^{\text{true}}$ ) on random regular social connections with average number of neighbours of  $k = 3, 4, 6$  respectively. Consistent with the results we have shown for lattices, the emergence of cooperation is promoted with misinformation in weak dilemmas and sparse networks, while  $f_c^{\text{mis}} < f_c^{\text{true}}$  can be observed as  $b$  increases and the networks become denser. (b) We obtain the theoretical stable frequency of cooperators (dashed line) through pair approximation on random regular social connections with  $k = 6$ . The theoretical (dashed line) and numerical results (solid line) with true information (black) and misinformation (orange) at population density  $d = 0.3, 0.5, 0.7$  are shown from left to right respectively. The total number of nodes of the random regular network is 2400.