Stimuli strategy and learning dynamics promote the wisdom of crowds

Li Zhenpeng[1,a] and Tang Xijin[2,b]

[1] School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, Zhejiang, China
[2] Academy of Mathematics and Systems Sciences Chinese Academy of Sciences, Beijing 100190, China

Abstract. Collective wisdom is the ability of a group to perform more effectively than any individual alone. Through an evolutionary game-theoretic model of collective prediction, we investigate the role that reinforcement learning stimulus may play the role in enhancing collective voting accuracy. And collective voting bias can be dismissed through self-reinforcing global cooperative learning. Numeric simulations suggest that the provided method can increase collective voting accuracy. We conclude that real-world systems might seek reward-based incentive mechanism as an alternative to surmount group decision error.

1 Introductions

Why and how to make the many are smarter than the few and how collective wisdom shapes business, economics, societies, and nations? Collective wisdom, also called group wisdom, and co-intelligence refer to shared knowledge arrived at by individuals and groups. Collective wisdom, a long-standing problem in the social, biological, and computational sciences has been recognized that group can be remarkably smart and knowledgeable when compared with the judgements of individuals [1].

Collective decisions are central to human societies, from small-scale social units such as families and committees to large-scale organization such as governments and international organizations. Currently, humanity is facing the most pressing collective decision challenges; for instance, global climate change negotiations, poverty reduction action, COVID-19 like global pandemics collaborative control, etc. Collective wisdom is closely related to the well-being of human society, since wicked group decision-making will bring negative effects to each individuals.

Indeed, both the animal world and humans are facing the same collective dilemma. The phenomena of collective decision-making have been found in many biological colonies, such as ants [2], bees [3], birds [4,5], and fishes [6].

Individuals sometimes exhibit collective wisdom, and other times, maladaptive herding is an enduring conundrum [7]. Many findings suggest fundamental cognitive benefits of grouping; meanwhile, there is also convincing arguments that interacting individuals may sometimes generate madness of crowds, and undermine

the collective wisdom [8]. For example, recent investigations demonstrate that herd behavior generated through social influence can undermine the wisdom of crowd effect in simple estimation tasks [9]. Herd behavior, an alignment of thoughts or behaviors of individuals in a group, or imitation without individual rational decision-making to be a cause of financial bubbles [10], and volatility in cultural markets [11,12].

There is a long-standing recognition, especially for humans, interdependence, and social influence between individual decisions may undermine the wisdom of the crowds effect. Collective wisdom optimization, in other words, with the aim to overcome the mob effect to the greatest extend and how to design an effective mechanism to regulate the wisdom and madness of interactive crowds should be the center of behavior research. Some pioneering studies have made some progress towards this focus. Such as maintaining individual diversity among group members [13], balancing individual and social information [14], increasing group size [15–17], improving the plasticity and feedback ability of interactive social network structure [18], and social influence [9,19]. Especially, recent empirical social learning strategies [7,20] with the aim to regulate the wisdom and madness of interactive crowds.

The essence of collective wisdom is individual-based group decision-making, which covers many subjects such as cognition, social psychology, organizational behavior, etc. In the famous work organization, Simon regards decision-maker as an independent management mode, and all the members of the organization are decision-makers who reasonably choose means to achieve a certain purpose. He believes that the essence of any organization is a complex model of information

[a] e-mail: lizhenpeng@amss.ac.cn (corresponding author)
[b] e-mail: xjtang@amss.ac.cn

communication and mutual relationship among group individuals [21].

Indeed, group decision-making is deeply rooted in cumulative bodies of social information [22], imitation [23], cooperation, trust and trustworthiness [24], and moral preference [25]. However, this question is challenging to study, and previous research has reached mixed conclusions, because collective decision outcomes depend on the insufficiently understood complex system of cognitive strategies, task properties, and social influence processes [26].

In particular, the group decision-making process cannot avoid the conflict between individual and collective interests or objectives. In such situation, collective wisdom is unavoidably to be riddled with social dilemmas [27] or tragedy of the commons, in which decisions that make sense to each individual can aggregate into outcomes in which everyone suffers. To understand the collective decision performance of social learners fully requires fine-grained quantitative studies of social learning strategies and their relations to collective dynamics, linked to sophisticated computational analysis.

This kind of research not only has the theoretical research value of social group psychology, but also has the practical application value of surmounting group non-cooperative behavior. For example, sustainability challenges often require collective action by groups with different background experiences and individual interests. To effectively initiate large-scale collective action to sustain natural resources in a world of global environmental change, how to design scientific and effective mechanism is a key problem. Indeed, exploring key conditions or principles for the emergence of collective intelligence have important implications from organization theory to economics, artificial intelligence, and cognitive science [28].

In the context of evolutionary games and the public cooperation, we point to a discovery of a general and widely applicable phenomenon; stimuli strategy [13] and learning dynamics are exploited for the effective resolution of madness of crowds. Specifically, with the enlightenment of learning-theoretic solution concepts for social dilemmas [20,29], we shift attention to a learning-theoretic alternative with the aim to resolve social dilemma, so as to improve the level of collective wisdom. And then as mechanism design in favor of our societies and social welfare, we hope to deep understand collective intelligence, so we can create and take advantage of the new possibilities it enables.

2 Ground truth-based collective voting

Lorenz et al. presented an definition of a collection wise based on the aggregate individual estimation if the estimated average of the population comes close to the true value [9]. In this case, single estimates are likely to lie far away from the truth, but the heterogeneity of numerous decision-makers generates a more accurate aggregate estimation. This consideration not only reflects the

nature of group decision-making, but also provides a train of thought and theoretic guide for our study.

In the following, we first present theoretic analysis on the conditions that the individuals needed to improve group voting accurate. We consider a binary outcome $Y \in \{+1, -1\}$, and $Y = -1$ is the ground truth which depends on many agents voting $x_i \in \{-1, +1\}, i = 1, ...n$. We model this outcome as an estimator of the average collective voting results of all attending individuals, that is

$$\hat{Y} = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{1}$$

At initial round $t = 1$, since, without any prior information, each agent has been assigned a subjective probability $q_{i,t}$ to choose ground truth -1, i.e., agent i prefers to choose -1 with probability $q_{i,t}$. It is assumed that the accuracy of an individual's prediction can be judged after each voting round. Then, in the next voting round, agents are motivated by a stimulus that offers rewards for making more accurate predictions through increasing probability to select ground truth. In the next section, we find that reinforcement learning dynamic stimulus mechanism really improves the accuracy of group decision-making through computer simulation.

Here, use two measures to assess collective intelligence accuracy. One measurement is self-reinforcing cooperative voting probability $Q_t = (q_{1,t}, ..., q_{n,t})$ of all agents for ground truth at round t. The other is the collective error at each round t, which is defined as the absolute distance between estimator $\hat{Y}_t$ that the collective vote agrees with the ground truth given the distribution $Q_t = (q_{1,t}, ..., q_{n,t})$ and ground truth Y. Equation (2) states that the collective error is equal to the average individual estimate minus the ground truth

$$e_t = |\hat{Y}_t - Y|. \tag{2}$$

Assuming each agent voting assessment $x_{i,t} \in \{+1, -1\}$ satisfies binary distribution with probability $p_{i,t}$ to choose $+1$, and probability $q_{i,t}$ prefer to select -1. The expected collective voting result is then given by Eq. (3)

$$E(\hat{Y}_t) = E\left(\frac{1}{n} \sum_{i=1}^{n} x_{i,t}\right) = \frac{\sum_{i=1}^{n} p_{i,t}}{n} - \frac{\sum_{i=1}^{n} q_{i,t}}{n}, \tag{3}$$

where $p_{i,t} + q_{i,t} = 1$. Suppose we design a feasible mechanism, it can enhance collective voting accuracy, such that for all voters, $q_{i,t} \to 1$ or $p_{i,t} \to 0$ when $t \to \infty$. Then, Eq. (3) gives us the theoretic result $E(\hat{Y}_\infty) = -1$. It means that collective intelligence is improved steadily as a function of voting round t.

For any voting round t, we prove that for any two voting results $x_{i,t}, x_{j,t}$, they are uncorrelated, since

$\mathrm{Cov}(x_{i,t}, x_{j,t}) = E(x_{i,t} x_{j,t}) - E(x_{i,t}) E(x_{j,t}) = (p_{i,t} - q_{i,t})(p_{j,t} - q_{j,t}) - (p_{i,t} - q_{i,t})(p_{j,t} - q_{j,t}) = 0$. Then, we have the variance of collective voting estimator

$$\mathrm{Var}(\hat{Y}_t) = \frac{1}{n^2} \left\{ \sum_{i=1}^{n} \mathrm{Var}(x_{i,t}) + 2 \sum_{i \neq j}^{n} \mathrm{Cov}(x_{i,t}, x_{j,t}) \right\} \tag{4}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(x_{i,t}) = \frac{1}{n} \{ 1 - (p_{i,t} - q_{i,t})^2 \}. \tag{5}$$

It gives that $\mathrm{Var}(\hat{Y}_t) \to 0$, as $t \to \infty$, if $q_{i,t} \to 1$, i.e., $p_{i,t} \to 0$.

Next, we shift attention to a self-reinforcing cooperative learning. We use computational experiments to assess the collective voting accuracy with the framework of stimuli strategy and learning dynamics.

3 Self-reinforcing cooperative learning

In this section, we illustrate that collective voting error can be dismissed through self-reinforcing global cooperative learning dynamics.

3.1 Learning dynamics in social dilemma

Social dilemmas mean that everyone enjoys the benefits of collective action, but free riders gain more without contributing to the common good. For decades, resolving or dismissing social dilemmas and promoting cooperation has been a fascinating focus from a broad range of disciplines. In social dilemma cases, everyone always receives a higher payoff for defecting than for cooperating, but all are better off if all cooperate than if all defect. Such situation can be formalized as 2-person games where each player can either cooperate or defect. For each player, the payoff for both cooperate R is greater than the payoff obtained for both defect is P, while one cooperates and the other defects, the cooperator obtains S, and the defector receives T. $R > T > P > S$ resulted in the game of Stag Hunt, fear but not greed. $T > R > S > P$ resulted in the game of Chicken, greed but not fear. And, $T > R > P > S$ resulted in the celebrated game of Prisoner's Dilemma (PD). The prisoner's dilemma is a representative example of non-zero sum game in game theory, which reflects that the individual's best choice is not the group's best choice. In other words, in a group, individuals make rational choices, but often lead to collective irrationality. As a classical example of an evolutionary game that is governed by pairwise interactions, Prisoner's Dilemma game model is more suitable for our research problems, i.e., conflict between individual and collective interests or objectives. Therefore, we choose Pris-

oner's Dilemma game instead of game of Chicken or Stag Hunt.

Macy and Flache [29] explored the learning dynamics that observed in 2-player 2-strategy social dilemma games. They used an elaboration of a conventional Bush–Mosteller stochastic learning algorithm [30] for binary choice cooperate or defect. In their model, the individuals learn from the previous interactions and outcomes and therefore adapt their decisions because of experience. Especially interesting is the case; they consider the influence of individual's aspiration, habituation level, and learning rate on 2-player no-static equilibria mixed strategy.

The stochastic learning algorithm is summarized as following. For each evolutionary game round, players decide whether to cooperate or defect at random. Each player's strategy is defined by the probability of undertaking each of the two actions available to them. After players have selected an action according to their probabilities, each one receives the corresponding payoff and revises her strategy. The revision of strategies takes place following a reinforcement learning approach: players increase their probability of undertaking a certain action if it leads to payoffs above their aspiration level, and decrease this probability otherwise. When learning, players in the Bush–Mosteller model use only information concerning their own past choices and payoffs, and ignore all the information regarding the payoffs and choices of their counterparts. However, in this paper, we focus on a different situation, players in our model use both information concerning their own past choices and all the information regarding the payoffs of others. The information of others is transmitted to individuals through average payoffs of all members. Meanwhile, our model follows the same learning mechanism as that in the Bush–Mosteller model. First, players (or voters in our model) calculate their stimulus for the action; second, they update their aspiration level; finally, both current round stimuli s_t and voting probability p_t result in the updated probability for her next round voting.

The above steps form a closed loop and identified as two no-static equilibria formally defined in [31]. One is named self-correcting equilibrium (SCE) which is characterized by dissatisficing behavior. The other is named the self-reinforcing equilibrium (SRE). An SRE corresponds to a pair of pure strategies (i.e., q_t subject to $\{0, 1\}$), such that its certain associated outcome gives a strictly positive stimulus to players. SRE can be interpreted as a mutually satisfactory outcome. The number of SRE in a social dilemma depends on rewarded players. Near an SRE, there is a high chance that the system will move towards it. Therefore our investigation is restricted to the analysis of SRE.

3.2 Description of the model

Our model is based on the Bush–Mosteller stochastic learning model, but we extend 2-person games to a n-voters collective voting situation. An learning dynamics of individuals cooperating to obtain a collective voting

benefit is here modeled as an evolution of cooperation, where the each voter interacts with other $n-1$ individuals via n-player Prisoner's Dilemma game. At each voting round, each individual voting result depends on the other $n-1$ voters.

The probability $q_{d,t+1}$ of taking a decision $d = -1$ at round $t+1$ for each voter is updated by the consequences of her previous decision $d \in \{+1, -1\}$ at time t. These consequences are described as negative (punishment) when voter choose $+1$, and positive (reward) stimulus s_t when voter choose ground truth -1. The outcome $\{(+1, +1), (-1, -1), (+1, -1), (-1, +1)\}$ induces a payoff $\Re_t$ corresponding to one of the value (R, P, T, S) for each voter. This payoff results in a corresponding stimulus for each agent associated with a voting decision $\{+1, -1\}$, such that described in Eq. (6)

$$s_{d,t} = \frac{\Pi_{d,t} - A_t}{\sup[|R - A_t|, |P - A_t|, |S - A_t|, |T - A_t|]}. \quad (6)$$

Often, people in social dilemmas attend more to the group's payoffs than to their own. Thereby, we set $\Pi_{d,t}$ is the average payoff rewarded for a voter i, given the other $n-1$ individuals payoff, i.e., $\Pi_{d,t} = \frac{1}{n-1} \sum_{j \neq i} \Re_{j,t}$, $j = 1, ..., n$. Then, we can compute and analyze the voting results and learning dynamics of the n individuals simultaneously. The denominator denotes the upper bound of possible difference between the mean payoff $\Pi_{d,t}$ for each individual and its aspiration level A_t, and therefore, $-1 \leq s_{d,t} \leq 1$. Then, the aspiration level is updated as

$$A_{t+1} = (1-h)A_t + h\Pi_{d,t}. \quad (7)$$

Aspiration level is a relevant aspect of decision-making [32]; the overall probabilities of voting result $\{+1, -1\}$ are relative to the aspiration level and expected utility. Equation (7) suggests that the updated aspiration level is a weighed mean between previous round aspiration and the averaged payoff with considering habituation parameter h. Where the habituation rate h describes the voter subjective will to change her previous round decision. For each voter, both current round stimuli $s_{d,t}$ and voting probability $q_{d,t}$ result in the updated probability as in Eqs. (8, 9) for her next round voting, with considering learning rate l. Such that in the case of $s_{d,t} \geq 0$

$$q_{d,t+1} = q_{d,t}(1 - ls_{d,t}), \quad (8)$$

for the case of $s_{d,t} < 0$,

$$q_{d,t+1} = q_{d,t} - (1 - q_{d,t})ls_{d,t}. \quad (9)$$

The learning rate l is a adjustable parameter between 0 and 1. Where $0 < q_{d,t+1} < 1$ and suggests a tendency for a agent to repeat rewarded voting result and to avoid punished decision. Although both habituation level and learning rate destabilize cooperation, the model implies

a cooperative SRE if (and only if) voters' aspiration levels are lower than the payoff for group cooperation. With this short sequence, the model is completely implemented and we can repeat this procedure for an arbitrary number of iteration.

4 Results and discussion

In our simulation, we focus on typical Prisoner's Dilemma game is formally characterized by a certain ordering of the four payoff values ($T_i = 4 > R_i = 3 > P_i = 1 > S_i = 0$), $i = 1, ..., n$. We set number of voters $n = 5000$, iterative learning round 1000. To model a more realistic situation, we consider the habituation rate and learning rate heterogeneous distribution among voters, and assume the initial parameters to be Gaussian distributed. Therefore, we set habituation rate vector $\overrightarrow{h} \sim \text{abs}(N(0.1, 0.04))$, learning rate vector $\overrightarrow{l} \sim \text{abs}(N(0.5, 0.05))$, and initial probabilities for voting ground truth vector subject to $\overrightarrow{q_1} \sim \text{abs}(N(0.5, 0.01))$. Where "abs" denotes absolute value operation; for example, $\text{abs}(N(0.1, 0.04))$ means that taking the absolute value of the random number from the normal distribution with expectation 0.1 and variance 0.04.

Aspiration yield unexpected insights into the global dynamics of cooperation in social dilemmas. Results from the two-person game suggest that a self-reinforcing cooperative equilibrium solution is viable only within a narrow range of aspiration levels. For this reason, we first analyze the aspiration levels influence on $P\{(-1, ..., -1)|SRE\}$ and mean of collective voting error. Where $P\{(-1, ..., -1)|SRE\}$ defined as the probability that $n = 5000$ voters select ground truth -1 on certain fixed aspiration level, and the mean of collective voting error $\overline{e}_a = \frac{1}{n} \sum_{t=1}^{1000} |Y - \hat{Y}_t|$ is defined as the average absolute error after each 1000 rounds collective voting for a fixed aspiration level a.

Under the assumption that all voters' aspirations are at the same level. To investigate the dependence of collective voting error $\overline{e}_a$ and $P\{(-1, ..., -1)|SRE\}$ on aspiration levels, for fixed initial parameter vector distribution $\overrightarrow{h} \sim \text{abs}(N(0.1, 0.04))$, $\overrightarrow{l} \sim \text{abs}(N(0.5, 0.05))$, $\overrightarrow{q_1} \sim \text{abs}(N(0.5, 0.01))$, we vary the aspiration level across the entire range of payoffs (from 0 to 4) in steps of 0.1, given they locked in a SRE within 1000 iterations. The corresponding results are shown in Fig. 1. We observed that $P\{(-1, ..., -1)|SRE\}$ decreases rapidly to zero with the increase of aspiration level a within $(0, 1)$. That means a fair interaction between individual and other members is only possible in a narrow range of aspiration levels, i.e., SRE solution is viable only within a narrow range of aspiration levels such as $(0, 1)$. Meanwhile, the mean of collective voting error $\overline{e}_a$ increases with the increase of aspiration levels when the level is greater than 1. The plots in Fig. 1 shows that to obtain acceptable collective voting
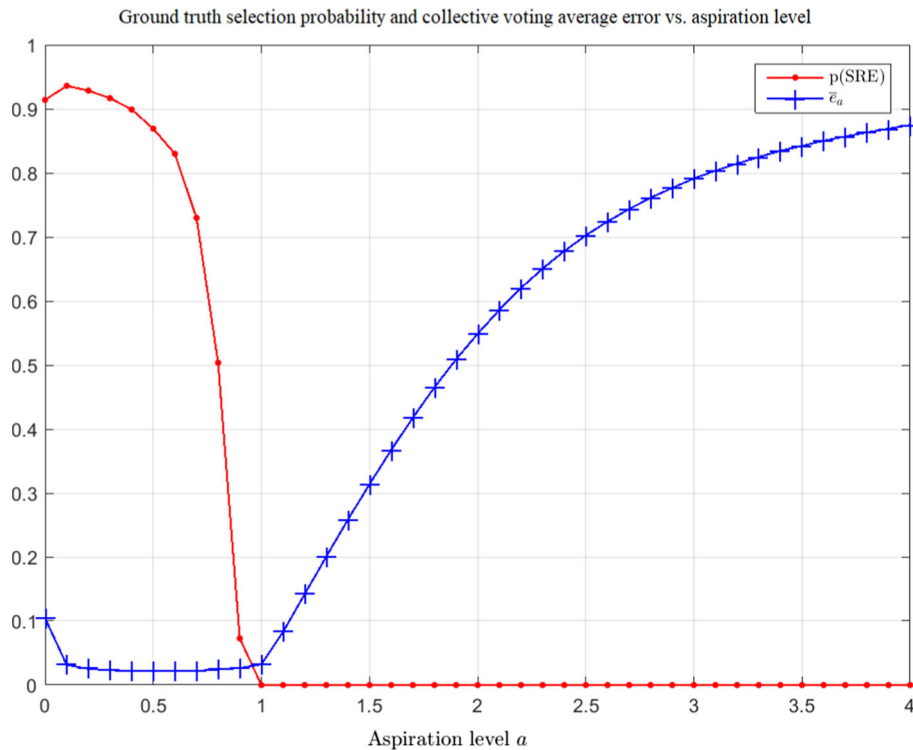
Fig. 1 Dependence of $P\{(-1,...,-1)|SRE\}$ and collective voting error $\bar{e}_a$ on the aspiration level a. It is the probability that voting for ground truth for 5000 voters, given they locked in a SRE within 1000 rounds. For all voters, we considered the fixed initial parameters distribution $\overrightarrow{h} \sim abs(N(0.1, 0.04)), \overrightarrow{l} \sim abs(N(0.5, 0.05)), \overrightarrow{q}_1 \sim abs(N(0.5, 0.01))$. Each plot is based on 100 replications at each level of aspiration

accuracy, we should choose an appropriate aspiration level.

In the following simulations, to restrict the aspiration level to the effective interval $(0,1)$, we set aspiration level $\overrightarrow{a} \sim U(0,1)$, where U denotes uniform distribution. Meanwhile, habituation rate $\overrightarrow{h} \sim$ abs$(N(0.1, 0.04))$, $\overrightarrow{l} \sim$ abs$(N(0.5, 0.05))$, and initial probabilities for voting ground truth $\overrightarrow{q}_1 \sim$ abs$(N(0.5, 0.01))$ remain unchanged.

The corresponding probabilities for voting ground truth $\overrightarrow{q}_t$ are shown in Fig. 2. Here, the probabilities $q_{i,t}$, $i = 1,...,5000$ are color-coded. To get an overview of the overall development of the population, for each iteration, the probabilities are sorted in ascending order and every row corresponds to an individual. As Fig. 2 shown after 50 rounds of reinforcement learning, nearly all voters prefer to select ground truth -1 with almost probability 1; we observe an obvious trustful SRE in the Prisoner's Dilemma.

We also calculate the average fraction of voting ground truth as a function of reinforcement learning step, given all players locked in a SRE within 1000 rounds for random initialized $\overrightarrow{h}$, $\overrightarrow{l}$ and viable aspiration level $\overrightarrow{a}$. We see that in Fig. 3, the ratio of voters with probability 1 to choose ground truth is evanescently large with increase of voting round.

At the same time, as we observed in Fig. 4, both the collective error e_t and $\mathrm{Var}(\hat{Y})$ tend to zero; the simula-
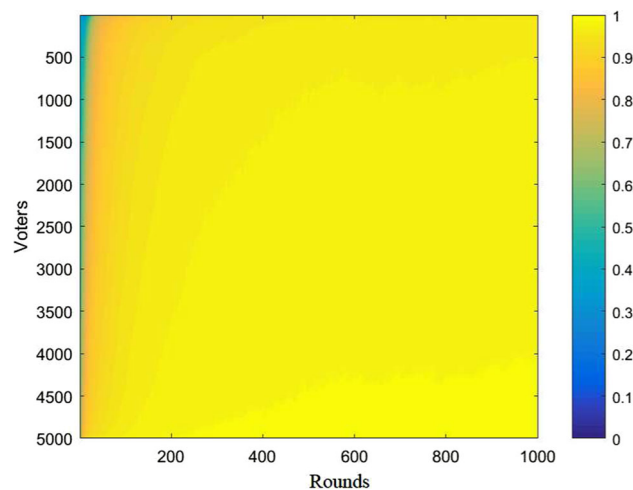

Fig. 2 The results show the development of the probability of choosing ground truth during 1000 rounds for each of the 5000 voters. Note that the probabilities in each column are sorted in ascending order, to visualize the development for the whole population. Here, we set the random initial parameters $\overrightarrow{h} \sim$ abs$(N(0.1, 0.04))$, $\overrightarrow{l} \sim$ abs$(N(0.5, 0.05))$, $\overrightarrow{q}_1 \sim$ abs$(N(0.5, 0.01))$, and $\overrightarrow{a} \sim U(0,1)$

tion results are well matched with our theoretic analysis. Furthermore, as the findings shown in Fig. 5, it can
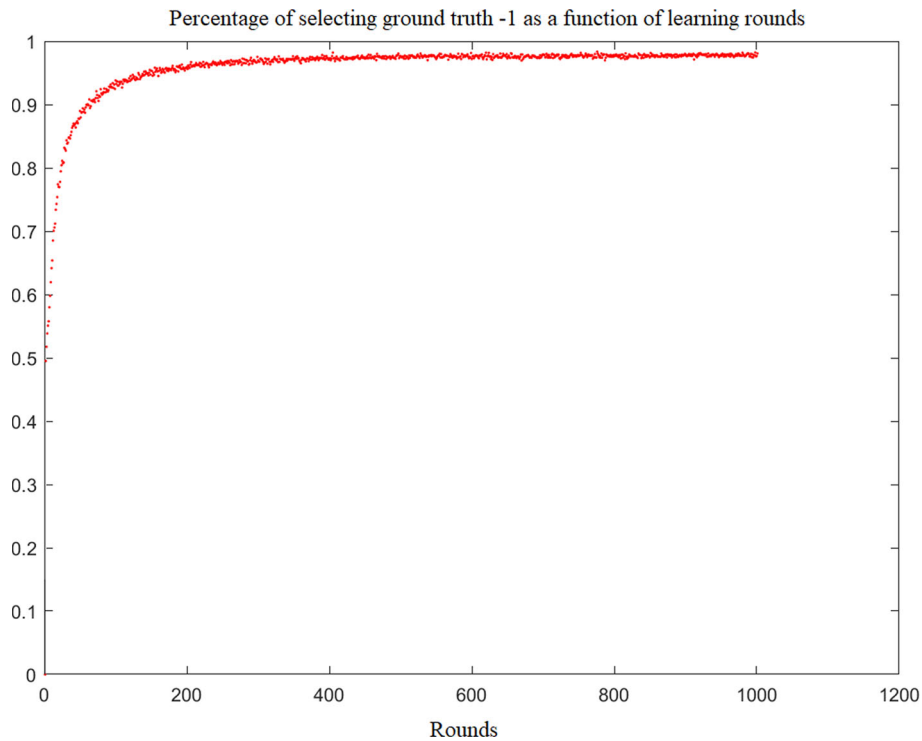
Fig. 3 Percentage of choosing ground truth as a function of learning rounds, given they locked in a SRE within 1000 iterations. For all voters, we considered the random initial parameters as that in Fig. 2
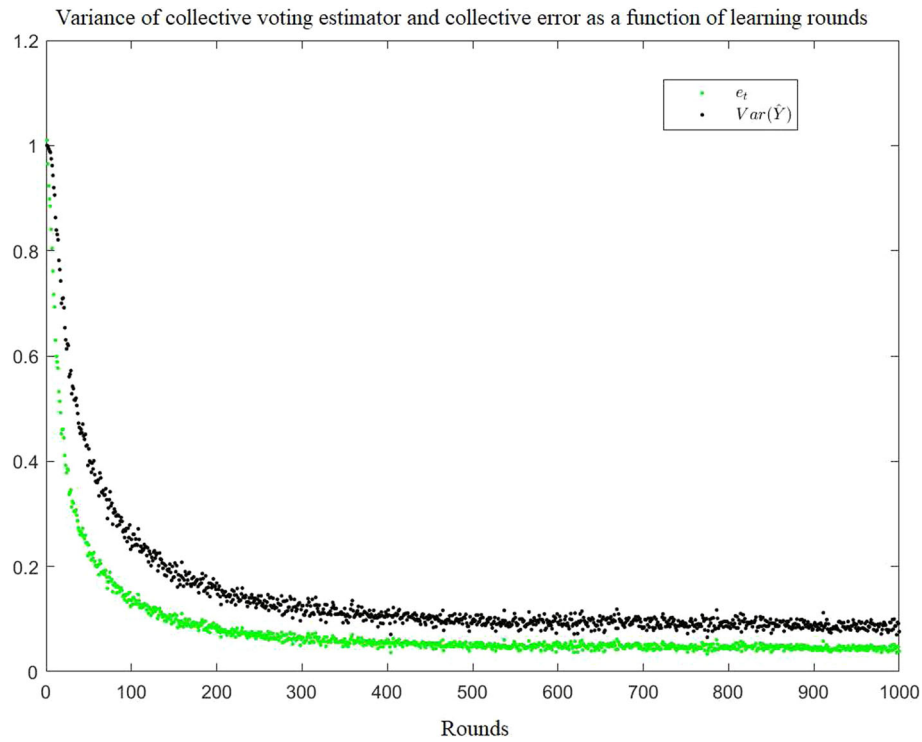
Fig. 4 Variation of collective voting error and the variance of collective voting estimator as a function of learning rounds

be find that $\hat{Y}$ is approximate to ground truth -1 as a function of learning rounds. Both theoretic analysis and numeric simulations suggest that reinforcement learning dynamic stimulus mechanism achieves its design performance in general, which really improves the accuracy of collective voting.
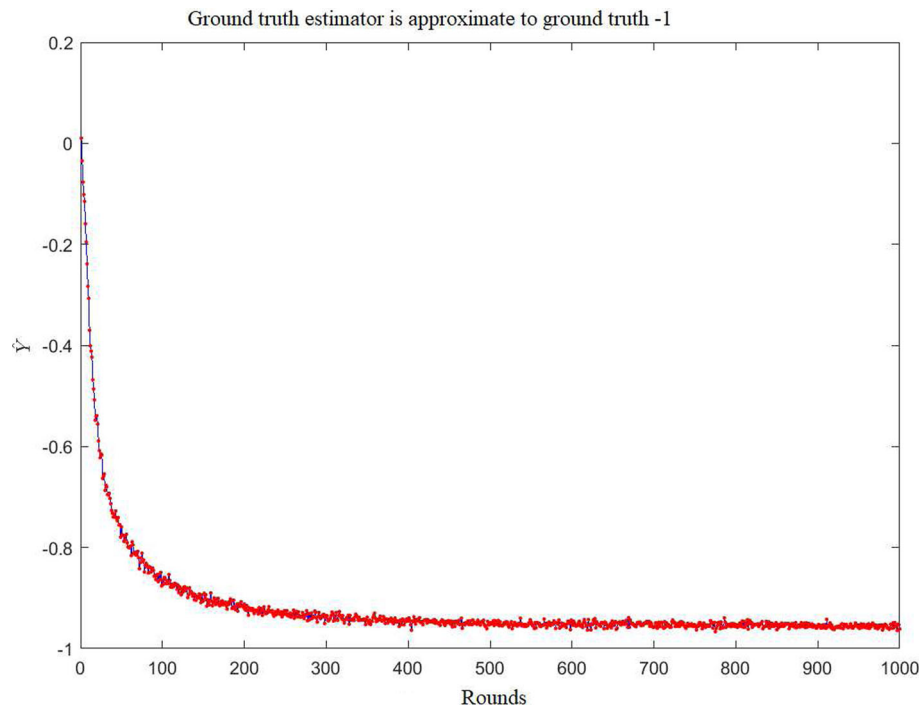
Fig. 5 Collective voting estimator $\hat{Y}_t$ is approximate to ground truth -1 as a function of collective voting learning rounds

Statistical explanations for our proposed swarm intelligence promotion approach argue that group accuracy relies on estimates taken from groups where individuals errors are correlated through stimuli strategy and learning dynamics mechanism. Thus, when the number of individuals in the group is large enough, although decentralized decision-making individuals may have estimates both far above and far below the true value, in aggregating these errors cancel out, leaving an accurate group judgement. The conclusion is agreed with the opinion of the classic wisdom of the crowds theory, increasing group size may increase collective accuracy [15–17]. However, we argue that increasing group size may not necessarily improve collective judgement accuracy without designing scientific and effective strategies to improve collective wisdom.

5 Conclusions

Collective wisdom is critical to solving many scientific, business, and other problems. In this paper, we show that collective intelligence can be improved through reinforcement learning incentive approach in evolutionary game. The study maps the landscape for the collective voting dynamics at the cognitive level, beginning with the simplest possible iterated social dilemma. The results convincingly support the fact that the evolution of cooperation promotes the wisdom of groups. Particularly, we show that reward-based incentive scheme might be a feasible alternative to overcome mob effect or herding effects and improve collective wisdom, and

shed light on the ability of groups to resolve collective action problems. We illustrate the usefulness of the framework to resolve social dilemma through stimuli strategy and learning dynamics, while promoting collective wisdom and enhancing the predictability or judgement of human social activity. Echoing the conclusion "wisdom of groups promotes cooperation in evolutionary social dilemmas" [28], our result suggests that cooperation in evolutionary social dilemmas also promotes "wisdom of groups".

Author contributions

All authors contributed to the conceptualization and research ideas. Li Zhenpeng designed the research plan and performed the simulation experiments. Li Zhenpeng contributed to reviewing and editing the paper. The authors declare no competing interests.

Availability of data and materials This manuscript has no associated data or the data will not be deposited. [Authors' comment : Since our simulation results are based on the average of multiple random initial results, different number of replications and model evolution steps may cause slights differences in the test results, eventhough the patterns in Figs. 1-5 are consistent. Therefore, we did not save the experimental data.]

References

1. F. Galton, Vox populi. Nature **75**, 1949 (1907)
2. J.E. Strassmann, D.C. Queller, Insect societies as divided organisms: the complexities of purpose and cross-purpose. Proc. Natl. Acad. Sci. USA **104**, S1 (2007)
3. K.M. Passino, T.D. Seeley, P.K. Visscher, Swarm cognition in honey bees. Behav. Ecol. Sociobiol. **62**, 3 (2008)
4. A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, M. Viale, Scale-free correlations in starling flocks. Proc. Natl. Acad. Sci. USA **107**, 26 (2010)
5. W. Bialek, A. Cavagna, I. Giardina, T. Mora, E. Silvestri, M. Viale, A.M. Walczak, Statistical mechanics for natural flocks of birds. Proc. Natl. Acad. Sci. USA **109**, 13 (2012)
6. A.J.W. Ward, J.E. Herbert-Read, D.J.T. Sumpter, J. Krause, Fast and accurate decisions through collective vigilance in fish shoals. Proc. Natl. Acad. Sci. USA **108**, 6 (2011)
7. W. Toyokawa, A. Whalen, K.N. Laland, Social learning strategies regulate the wisdom and madness of interactive crowds. Nat. Hum. Behav. **3**, 2 (2019)
8. J. Surowiecki, The wisdom of crowds: Why the many are smarter than the few. Abacus 295, (2004)
9. J. Lorenz, H. Rauhut, F. Schweitzer, D. Helbing, How social influence can undermine the wisdom of crowd effect. Proc. Natl. Acad. Sci. USA **108**, 22 (2011)
10. I. Yousaf, S. Ali, S. Shah, Herding behavior in Ramadan and financial crises: the case of the Pakistani stock market. Financial Innov. **4**, 1 (2018)
11. M.J. Salganik, P.S. Dodds, D.J. Watts, Experimental study of inequality and unpredictability in an artificial cultural market. Science **311**, 5762 (2006)
12. L. Muchnik, S. Aral, S.J. Taylor, Social influence bias: a randomized experiment. Science **341**, 6146 (2013)
13. R.P. Mann, D. Helbing, Optimal incentives for collective intelligence. Proc. Natl. Acad. Sci. USA **114**, 20 (2017)
14. C. List, C. Elsholtz, T.D. Seeley, Independence and interdependence in collective decision making: an agent-based model of nest-site choice by honeybee swarms. Philos. Trans. Roy. Soc. Lond. Ser. B Biol. Sci. **364**, 1518 (2009)
15. C. List, Democracy in animal groups: a political science perspective. Trends Ecol. Evolut. **19**, 4 (2004)
16. A.J. King, G. Cowlishaw, When to use social information: the advantage of large group size in individual decision making. Biol. Lett. **3**, 2 (2007)
17. M. Wolf, R. Kurvers, A. Ward et al., Accurate decisions in an uncertain world: collective cognition increases true positives while decreasing false positives. Proc. Biol. **280**, 1756 (2013)
18. A. Almaatouq, A. Noriega-Campero, A. Alotaibi, P.M. Krafft, M. Moussaid, A. Pentland, Adaptive social networks promote the wisdom of crowds. Proc. Natl. Acad. Sci. USA **117**, 21 (2020)
19. J. Becker, D. Brackbill, D. Centola, Network dynamics of social influence in the wisdom of crowds. Proc. Natl. Acad. Sci. USA **114**, 26 (2017)
20. D. Jia, H. Guo, Z. Song, L. Shi, X. Deng, M. Perc, Z. Wang, Local and global stimuli in reinforcement learning. New J. Phys. **23**, 8 (2021)
21. H. Simon, J.G. March, *Organization* (John Wiley Press, New York, 1958)
22. H. Wang, B. Wellman, Social connectivity in America: changes in adult friendship network size from 2002 to 2007. Am. Behav. Sci. **53**, 8 (2010)
23. M.G. Zimmermann, V.M. Eguíluz, M. SanMiguel, Coevolution of dynamical states and interactions in dynamic networks. Phys. Rev. E. **69**, 62 (2004)
24. A. Kumar, V. Capraro, M. Perc, The evolution of trust and trustworthiness. J. R. Soc. Interface **17**, 169 (2020)
25. V. Capraro, M. Perc, Mathematical foundations of moral preferences. J. R. Soc. Interface **18**, 175 (2021)
26. V. C. Yang, M. Galesic, H. McGuinness, A. Harutyunyan, Dynamical system model predicts when social learners impair collective performance. Proc. Natl. Acad. Sci. USA 118 (35) 2021
27. R.M. Dawes, Social dilemmas. Annu. Rev. Psychol. **31**, 1 (1980)
28. A. Szolnoki, Z. Wang, M. Perc, Wisdom of groups promotes cooperation in evolutionary social dilemmas. Sci. Rep. **2**, 576 (2012)
29. M. W. Macy, A. Flache, Learning dynamics in social dilemmas. Proc. Natl. Acad. Sci. USA 99, suppl 3, 2002
30. M. Iosifescu, R. Theodorescu, On bush-mosteller stochastic models for learning. J. Math. Psychol. **2**, 1 (1965)
31. S.S. Izquierdo, L.R. Izquierdo, N.M. Gotts, Reinforcement learning dynamics in social dilemmas. J. Artif. Soc. Soc. Simul. **11**, 2 (2008)
32. E. Diecidue, J. van de Ven, Aspiration level, probability of success and failure, and expected utility. Int. Econ. Rev. **49**, 2 (2008)