Discrimination emerging through spontaneous symmetry breaking in a spatial prisoner's dilemma model with multiple labels

Gorm Gruner Jensen ®,[*] Frederik Tischel ®, and Stefan Bornholdt
Institute for Theoretical Physics, University of Bremen, 28359 Bremen, Germany

Social discrimination seems to be a persistent phenomenon in many cultures. It is important to understand the mechanisms that lead people to judge others by the group to which they belong rather than individual qualities. It was recently shown that evolutionary (imitation) dynamics can lead to a hierarchical discrimination between agents marked with observable, but otherwise meaningless, labels. These findings suggest that it can give useful insight to describe the phenomenon of social discrimination in terms of spontaneous symmetry breaking. The investigations so far have, however, only considered binary labels. In this contribution we extend the investigations to models with up to seven different labels. We find the features known from the binary label model remain remarkably robust when the number of labels is increased. We also discover a new feature, namely that it is more likely for neighbors to have strategies which are similar, in the sense that they agree on how to act toward a subset of the labels.

I. INTRODUCTION

Discrimination is often defined as treating individuals differently because of the groups to which they belong rather than individual traits (or qualities). Most of the empirically observed discrimination can be explained by ingroup favoritism [1,2]. That is, people tend to act in favor of those who are similar to themselves at the expense of those who are different. An increasing body of research, however, is suggesting that discrimination cannot be explained by ingroup favoritism alone. A number of experiments have shown that many intergroup relations are asymmetric in the sense that members of one of the groups show much less, sometimes even negative, ingroup favoritism than members of the other [3–8]. These findings seem to indicate that there exists a sort of hierarchy of social status between different groups. In this paper we will explore a minimalistic evolutionary game theory model in which persistent hierarchical discrimination can emerge through spontaneous symmetry breaking.

Most of the evolutionary game theory literature has been a search for mechanisms which promote cooperative behavior through evolutionary dynamics [9,10]. One mechanism which has been thoroughly studied is the so-called tag-based cooperation [11–19]. This is of special interest to us, because the introduction of observable tags makes it possible to define discrimination in a very simple way. One can say that an agent's behavior—or strategy—is discriminating if it is different toward peers who have different tags but identical behaviors. Most of the models of tag-based cooperation are, however, not directly applicable for describing persistent hierarchical discrimination. Some papers have already been very explicit about the close relation between tag-based cooperation and

ingroup favoritism [17,20]. Also, the behavior of the tag-based cooperation models tends to express cyclic or wavelike dynamics, with a constant renewal of the dominating tags and strategies known as the "chromodynamics of cooperation" [13].

Recently, it was demonstrated that persistent hierarchical discrimination can emerge if the evolutionary dynamics only works on strategies but not on the labels [21]. By starting from a spatially extended system—another well-known mechanism for promoting cooperation [22–24]—it was shown that under high selection pressure the cooperation would partially fail in a way which breaks the symmetry between two groups of agents distinguished only by an otherwise meaningless label. It was shown that the proposed dynamics consistently leads to a state where, dependent on parameters, either the minority or the majority is systematically favored. Human societies, however, consist of more than two types of people. Humans can, for example, have many different religions, countries of origin, eye colors, and so on. It is, therefore, natural to ask whether the same mechanism can also lead to discrimination if there are more than two different labels in the system.

In the literature on tag-based cooperation, it is quite common to define a model which allows for a large variation of different tags, and some recent studies have been addressing how phenotypic variation arises in the first place [18,19]. In this paper we will examine an extension of the hierarchical discrimination model [21] with up to seven different labels. This leads to a richer set of possible model outcomes, as the number of competing strategies grows exponentially with the number of labels. We will show that much of the original structure is preserved, in particular that the number of labels have almost no impact on the parameter regions dominated by unconditional cooperation or defection. This is a remarkable result considering that the fraction of nondiscriminating strategies decreases exponentially.

————
[*]ggjensen@itp.uni-bremen.de

II. MODEL

Consider a population of N agents occupying the nodes of a graph. The graph edges represent an agent's neighborhood. Agents interact with their neighbors in a prisoner's dilemma type game where they can either cooperate or defect. Each agent has one of L distinct labels which can be observed by its neighbors. The label is the only observable difference between agents. An agent may therefore discriminate by cooperating with those neighbors who carry some labels while defecting against those carrying others. A strategy specifies whether to cooperate with or defect against neighbors with each of the L labels. There are 2^L different possible strategies since a strategy can be represented by one binary variable for each label.

Cooperation costs an amount c for the donating agent and gives a benefit b to the receiver. The payoff of an agent is calculated as the sum of the benefits the agent receives minus the sum of the costs the agent pays:

$$p_i = \sum_{j \in \mathcal{N}_i} b\sigma_j(\lambda_i) - c\sigma_i(\lambda_j), \qquad (1)$$

where $\mathcal{N}_i$ is the set of neighbors of agent i, λ_i is the label of agent i, and σ_i is a function describing its strategy. $\sigma_i(\lambda_j) = 1$ if the strategy is to cooperate with the label of agent j and $\sigma_i(\lambda_j) = 0$ if it is to defect. It should be noted that the payoff is a simple function of the state—the labels and strategies of the agent and its neighbors—and that it is not accumulated over time.

The dynamical variables in our model are the agents' strategies. These change according to the following rule: First, we choose a random agent with uniform probability whose strategy will be updated. With a small probability μ the agent will "mutate," i.e., choose a new strategy at random with uniform probability distribution. Most of the time, with probability $1 - \mu$, the selected agent will copy the strategy of one of its neighbors. That neighbor is chosen with a probability proportional to its "fitness." The fitness of agent i, f_i, is related to its payoff p_i via the expression $f_i = \exp(wp_i)$. Here w is a global parameter which we will refer to as the selection pressure. When the selection pressure is very small, $w \to 0$, it is almost equally likely to choose any neighbor independent of their payoff. When the selection pressure is large, $w \to \infty$, the neighbor with the largest payoff will almost certainly be chosen.

Notice that when an agent chooses a new strategy it is indifferent to how the strategy matches with the labels. For example, there is nothing to hinder that an agent with a blue label copies a strategy from a neighbor with a green label which dictates only to cooperate with green neighbors.

In the case where there are $L = 2$ different labels, our model is the same as that studied in [21]. In the case where there is only $L = 1$ type of agents, it is a slight variation of a model presented in 2005 by Ohtsuki *et al.* [23], designed to demonstrate that evolutionary dynamics can promote cooperation in systems with spatial structure. Our model varies from the one proposed by Ohtsuki *et al.* by using the exponential function $[f = \exp(wp)]$ in the relation between fitness and payoff rather than an affine function $(f = 1 - w + wp)$. The two functions converge in the limit of vanishing selection pressure $(w \to 0)$.

III. RESULTS

The main results presented in this paper are all obtained from a system with 100×100 agents arranged in a regular square lattice with periodic boundary conditions. The labels are randomly assigned to each agent with independent uniform probability. To reduce the number of parameters we have used a small constant mutation rate of $\mu = 0.001$. The nonzero mutation rate prevents the system from getting stuck in absorbing single-strategy fixed points. One benefit of this is that we can start every simulation in the simplest strategy configuration in which no agent cooperates with anyone. We have also explored the model with agents arranged in both random regular graphs and Erdős-Rényi graphs. These results are collected in section A of the Supplemental Material [25] for comparison.

A. Three labels

To build an understanding of how our model behaves, we start by presenting five examples of typical dynamics arising when agents are arranged in a square lattice and the number of distinct labels is $L = 3$ (see Fig. 1). These examples illustrate the variety of behaviors observed at different values of cooperation benefit b and selection pressure w. Each example consists of two subfigures: One time series of the fraction of the population following each of the eight possible strategies and one snapshot showing how the strategies are distributed on the lattice at the end of the time series. Because of the nonzero mutation rate it is impossible for the model to have any fixed points, but the time series clearly indicate that the systems tend to reach an attractor in which the fractions of agents following each strategy fluctuate with small variations around some constant value. In the three examples A, B, and D the stationary distributions are clearly dominated by a single strategy. We find that this is a general result as long as the selection pressure w is small ($w \lesssim 0.2$) and the dynamics run long enough. This is reminiscent of the absorbing "single-strategy" states in the voter model [26,27]. In example C the system gets separated into two large "patches", each dominated by a single strategy. This type of state can remain for a very long time, but once one of the strategies has displaced the other, it is extremely unlikely for the other to return. In that sense the dynamics between strategies favoring the same number of labels appears to resemble that of voter models with some amount of surface tension [28–30]. When the selection pressure is large, such as in example E, we find that the lattice tends to fracture into many small patches. This situation is more complicated, and we will return to a more thorough analysis later.

Each simulation was initiated with the strategy configuration in which all agents are "not cooperating with anyone." The parameters chosen for each of the five examples are marked in the phase diagram in Fig. 2.

A: Our first example has a low cooperation benefit $b = 2$ and intermediate selection pressure $w = 0.1$. At these parameters, the system stays dominated by the strategy C(none) (not cooperating with anyone).

B: With a slightly higher cooperation benefit $b = 3$, the system ends up with a majority of agents who cooperate with
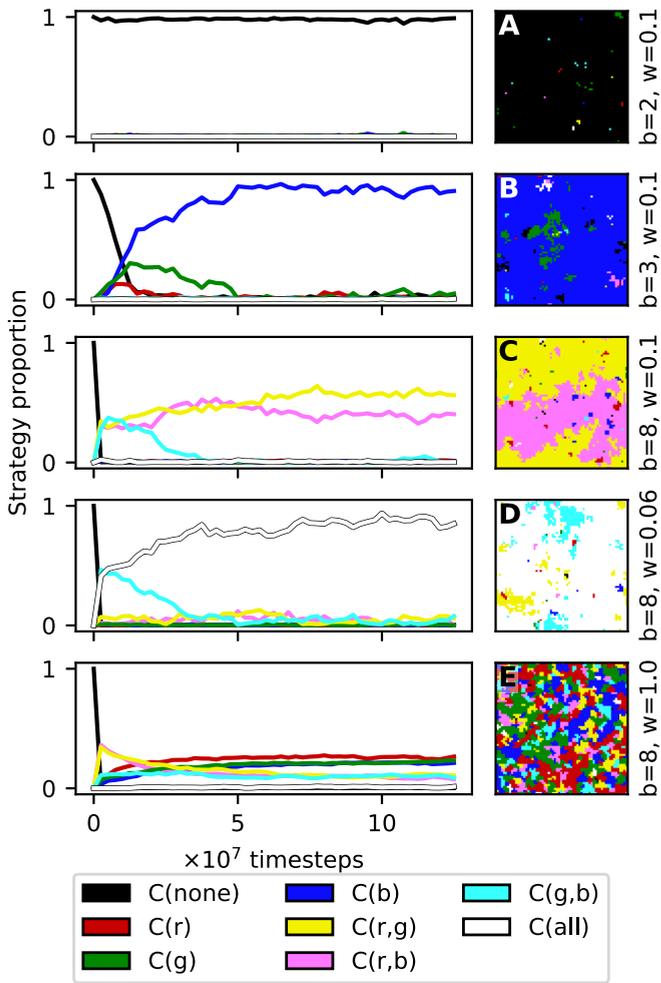
FIG. 1. Left: Time series of the strategy proportion for squared lattices with periodic boundary conditions, 10 000 agents and three labels (red, green, and blue) for different benefits b and selection pressures w with constant cost $c = 1$ and constant mutation rate $\mu = 0.001$. Starting with anyone having the strategy to not cooperate. Right: Snapshots of the strategy distributions at the end of the corresponding time series. Each color represents a different strategy, e.g., red represents the strategy of cooperating with neighbors with red labels only, $C(r)$; yellow represents the strategy of cooperating with neighbors if their labels are red or green, $C(r, g)$; and so on. The chosen values of b and w correspond to the different phases shown in phase diagram in Fig. 2.

one of the three labels. In our example it is the agents with green labels who receive the positive treatment by almost everyone, but since the labels are defined symmetrically, the dominant strategy could just as well have been only cooperating with red, $C(r)$, or with green, $C(g)$. We say that these strategies belong to strategy class 1—or $C(1)$—because they single out one of the labels as the only receiver of cooperation. Looking at the time series we see that all 3 $C(1)$ strategies are initially expanding by out-competing the $C(\text{none})$ strategy. However, when there are no more complete defectors to displace one of the $C(1)$ strategies ends up suppressing the others and eventually dominates the entire population. This is an example of spontaneous symmetry breaking, as the local dynamics result in a systemwide difference between

the outcome for agents with different labels, even though the labels are symmetrically defined.

C: The dynamics at high cooperation benefit, $b = 8$, are very similar to those described in example B, except that here it is the $C(2)$ strategies which are dominating the system. These are the strategies which cooperate with two of the label, e.g., cooperate with red and blue neighbors, but not those with green labels. Again we see that the system ends up being dominated by just one of these strategies. It is worth noting that in the very early rounds, while most of the agents still have the strategy not to cooperate with anyone, there is a brief rise of agents cooperating with everyone, $C(\text{all})$.

D: If, compared to example C, the selection pressure is a little smaller, $w = 0.06$, then the balance switches between the strategies $C(2)$ and $C(\text{all})$. Here the system ends up in a state where the large majority of agents are cooperating with all of their neighbors, independently of their label.

E: If, however, the selection pressure is increased to a high value, $w = 1$, then the system fractures into small patches each dominated by a single strategy. Globally, this results in a mix of all strategies of strategy classes 1 and 2, with strategy class 1 being more dominant. A number of experiments have shown that many intergroup relations are asymmetric in the sense that members of one of the groups show much less, sometimes even negative, ingroup favoritism than members of the other [3–8]. The nondiscriminating strategies are almost nonexisting, so the label symmetry is broken locally almost everywhere. However, the symmetry is restored at the macroscopic (systemwide) scale, because the patches of different strategies from the same strategy class occur with the same frequency and size distribution.

The examples above are representative of qualitatively different model behaviors. In the next section we will focus on how the behavior depend on the model parameters, by looking at how examples are arranged in parameter space. The left panel of Fig. 2 is a phase diagram showing which strategy class is the most dominating as a function of the cooperation benefit b and the selection pressure w. One can clearly distinguish four of the five phases exemplified in Fig. 1. While the distinction between B and E were clear in the examples, it cannot be seen by looking at this measure alone.

The two panels in the middle show parameter scans made at fixed cooperation benefits, $b = 8.0$ and $b = 2.0$, respectively, as indicated by the horizontal lines in the phase diagram. These show what fractions of the population belong to each strategy class as a function of the selection pressure w. At $b = 8$ we observe two transitions between different phases: One is a sharp transition at $w \approx 0.05$ from the phase where everyone cooperates with everyone else to the phase where almost all agents has a strategy of class 2. The other is a smooth transition at $w \approx 0.15$ to the state where strategy class 1 and 2 coexist with class 1 being more common. For $b = 2$ we see just one sharp transition from the phase where nearly nobody cooperates with anyone to a phase strongly dominated by strategy class 1.

The two panels on the right show parameter scans similar to those in the middle, but with varying cooperation benefit b and fixed selection pressures, $w = 0.1$ and $w = 1.0$, as marked by the vertical lines in the phase diagram. At $w = 0.1$
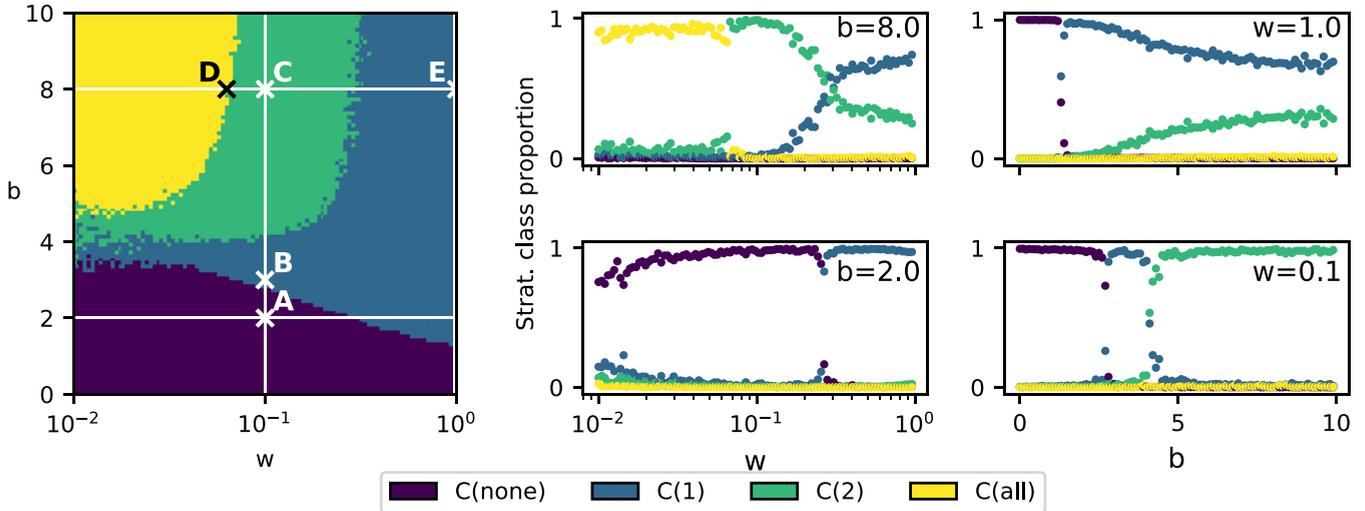
FIG. 2. Left: Phase diagram. The color of each pixel correspond to the most frequent strategy class after 1.25×10^8 time steps in a single simulation with the corresponding selection pressure w and cooperation benefit b. The parameters marked A–E correspond to the five examples shown in Fig. 1. Middle: Parameter scans showing strategy class proportion as a function of selection pressure w for fixed cooperation benefit b (top: $b = 8.0$, bottom: $b = 2.0$) corresponding to the horizontal lines in the phase diagram. Right: Parameter scans showing strategy class proportion as a function of cooperation benefit b for fixed selection pressure w (top: $w = 1.0$, bottom: $w = 0.1$), corresponding to the vertical lines in the phase diagram. The results are obtained on 100×100 squared lattices with periodic boundary conditions, in a system with $L = 3$ different labels, at a constant mutation rate $\mu = 0.001$ and a constant cost of cooperation $c = 1$.

we observe three different phases separated by two sharp transition. The transition between almost no cooperation and the phase dominated by strategy class 1 happens at $b \approx 2.8$. The phases dominated by strategy classes 1 and 2, respectively, are separated by a transition at $b \approx 4.5$ which is almost as sharp. At $w = 1.0$ there is a sharp transition at $b \approx 1.3$ between the phase with almost no cooperation to one where almost all agents have a strategy from class 1. When the cooperation benefit b is increased above this transition, we observe a smooth change where an increasing fraction of the agents ends up with strategies from class 2, resulting in the "mixed phase" as illustrated by example E in Fig. 1.

Based on these observations we can say that for a population with three different labels and agents arranged on a squared lattice there are sharp transitions between the phases represented by the examples A–D in Fig. 1 and smooth transition from B to E and from C to E.

B. More labels

To learn how the model behaves in systems with $L > 3$ labels, Fig. 3 shows parameter scans for systems with up to seven different labels. The plots show the most common strategy class as a function of cooperation benefit b and selection pressure w. We see that the regions of parameter space in which the system is dominated by unconditional cooperation or defection are almost unaffected when changing the number of labels. As we increase the number of labels, the number of strategy classes goes up as well. Consequently, the discriminating parameter region is subdivided into an increasing number of subregions, characterized by which strategy class ends up dominating the system. This division follows a very simple structure. Nearest to the region dominated by unconditional defection, the system will be dominated by

a strategy which cooperates with one of the labels (e.g., cooperate only with blue). When going a little further, the dominating strategies are those who cooperate with two of the labels (e.g.. cooperate with blue and green but not with the rest), and so on. Nearest to the region of unconditional cooperation we find the system dominated by strategies which discriminate negatively against a single label (e.g., cooperate with all neighbors except the blue).

The phase diagrams in Fig. 3 only show which strategy class ends up as the most abundant at a given parameter set. Therefore, it hides the finer structures, such as whether the system ends up being dominated by a single representative of the strategy class or if the stationary state is a mixture of many strategies as, for example, in example E in Fig. 1. In the $L = 3$ case, one can get a good sense of this transition by looking at the one-dimensional parameter scans shown in Fig. 2 and paying attention to how big a fraction of the agents follow the most abundant strategy class. Similar figures, but with up to $L = 7$ different labels, can be found in the Supplemental Material section A [25]. They indicate that the parameter region in which the system is characterized by fractured coexistence (similar to example E) does not depend on the number of labels.

In summary, the phase-diagram capturing the model's long-term behavior seems to be remarkably independent of the number of labels.

C. Further analysis

The results presented so far have all been limited to the case where agents have four neighbors. In this next section we will characterize how the system changes when we change the connectivity. Figure 4 shows simulation results from systems with three labels and agents arranged in grids with three labels
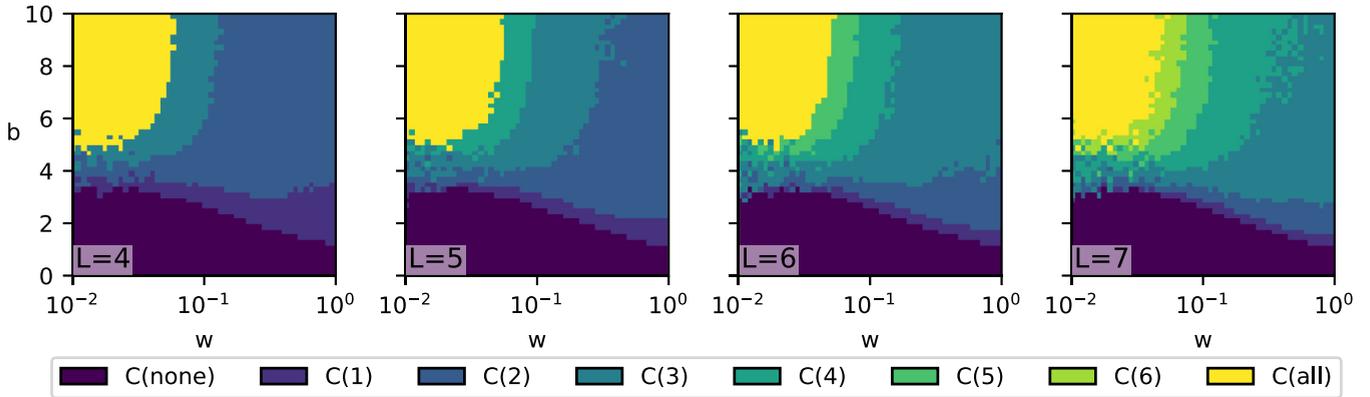
FIG. 3. Phase diagrams for systems with 4, 5, 6, and 7 possible labels (counting from left to right). The color of each pixel correspond to the most frequent strategy class after 1.25×10^8 time steps in a single simulation with the corresponding selection pressure w and cooperation benefit b. The results are obtained on 100×100 squared lattices with periodic boundary conditions, in systems with up to $L = 7$ different labels, at a constant mutation rate $\mu = 0.001$, and a constant cost of cooperation $c = 1$.

and up to 12 neighbors. Interested readers can find a detailed description of the lattice structures in the Supplemental Material section B [25].

Figure 4(a) shows three phase diagrams—similar to those presented in the previous figures—with 4, 8, and 12 neighbors, respectively. It can be seen that the cooperation benefit b have to be kept proportional to the connectivity k to get similar results. This was expected based on the results from Ohtsuki *et al.* [23], who said that cooperation occurs if the benefit-to-cost ratio exceeds the connectivity. It can also be seen that the system's stationary state is almost independent of cooperation benefits, when this is above ~$2k$ (marked by the horizontal lines). This means that when the cooperation benefit is high, the transition point between phases of unconditional cooperation and discrimination is almost only controlled by the selection pressure and the connectivity. It is clear to see that when the connectivity is higher, the transition into a discriminating phase happens at lower selection pressure. A more detailed visualization of this can be seen in Fig. 4(b). Here we compare parameter scans at varying the selection pressure w for connectivities ranging from $k = 3$ to $k = 12$ and fixed cooperation benefit $b = 2k$. It can be seen how reduction of the selection pressure needed to increase discrimination has a slightly convex course.

Figure 4(c) gives some more detail about how the strategy distribution depends on connectivity at k when both the selection pressure and cooperation benefit are high ($w = 1$ and $b = 2k$). The round dots show what fraction of the population ends up in each strategy class, as measured on the left axis. It can be seen that when the connectivity is higher, the gap between the strategy classes 1 and 2 becomes bigger. When $k \gtrsim 10$, almost all agents end up with a $C(1)$ strategy, even though all other parameters are the same as in example E in Fig. 1 where the stationary distribution was a disorderly mixture of small patches with different strategies. To further quantify this difference, the red crosses show the fraction of neighbors who have different strategies (boundary fraction). This is measured as the number of links connecting agents with different strategies divided by the total number of links. The "boundary fraction" is a lot smaller when the connectivity is higher, as the system goes

toward a state dominated by a single strategy. This indicates that increasing the connectivity pushes the smooth transition between the phases represented by the examples E and B in Fig. 2 toward higher selection pressure. To give a more intuitive visualization of this difference, Fig. 4(d) shows three snapshots of how the stationary strategy distribution look for three different connectivities—$k = 4$, $k = 8$, and $k = 12$, respectively.

It is not immediately obvious, but a careful study of snapshots like the one in Fig. 4 with connectivity $k = 4$ or example E in Fig. 1 reveals an interesting pattern in how frequently different strategies occupy neighboring nodes. Two different strategies are more likely to occupy neighboring nodes if they agree on their behavior towards more of the labels, i.e. their Hamming distance is small. e.g., a patch with the yellow strategy (cooperate with green and red) share, on average, significantly more border with the green or red patches than with patches of the blue strategy. In order to quantify this observation we can draw graphs like the one in Fig. 5, where each node represents one of the discriminating strategies, and the thickness of the link between two nodes is proportional to the number of neighbor pairs with one neighbor following each of the corresponding strategies (on average over an ensemble of individual simulations). The result varies with the choice of parameters. Here we have chosen to emphasize an example at cooperation benefit $b = 8$ and selection pressure $w = 10^{-0.54}$. This point is near the transition between the phases dominated by $C(1)$ and $C(2)$ strategies, so all the discriminating strategies are approximately equally abundant. We have chosen not to include the unconditional strategies, since they are almost nonexistent at these parameters. We have also chosen not to visualize the self-links, because these would be many times stronger than the links between different strategies. The figure shows that strategies are more likely to occupy neighboring when they agree on how to behave toward more of the labels. A handful of other examples are included in the Supplemental Material section C [25]. These show that the result is qualitatively robust at for a wide range of parameters near the transition between phases where different strategy classes coexist, except for very low values of selection pressure where the signal drowns in noise.

FIG. 5. Left panel: Graphs representation visualizing the strategy distributions. Each node represents one of the discriminating strategies. The thickness of a link between two nodes is proportional to the fraction of neighbor pairs with one agent following each of the corresponding strategies after 10^8 time steps averaged over 100 simulations with the following parameters: Each agent is connected to the $k = 4$ nearest neighbors in a 100×100 square lattice and has one of $L = 3$ possible labels chosen independently with uniform probability. We used cooperating benefit $b = 8$, selection pressure $w = 10^{-0.54}$, and mutation rate $\mu = 0.001$, and we initiated all simulations with all agents following the "defect all" strategy. Right panel: Example of one of the 100 data points combined to make the graph. Each color correspond to a different strategy in the same way as in Fig. 1.

FIG. 4. (a) Two-dimensional parameter scans similar to the one shown in Fig. 2 but with different for connectivities $k = 4$, 8, and 12 (from left to right). (b) Parameter scans showing the most common strategy class as a function of selection pressure w and connectivity k for benefit $b = 2k$. (c) Round dots show the fraction of each strategy class in the stationary state for different lattice-connectivities k. Red crosses show the fraction of neighbor pairs with different strategies. The simulations use cooperation benefit $b = 2k$ and a very high selection pressure $w = 1.0$. (d) Strategy distributions for connectivity $k = 4$, 8, and 12 (from left to right), benefit $b = 2k$, and selection pressure $w = 1$. The color-coding is the same as in Fig. 1. The results are obtained on lattices with 100 agents and periodic boundary conditions. The number of different labels is $L = 3$, and the mutation rate and the cost of cooperation are both constant, $\mu = 0.001$ and $c = 1$, respectively. The colors in panels (a)–(c) correspond to the different strategy classes. Each colors in panels (d) correspond to a different strategy in the same way as in Fig. 1.

While all the results presented in the so far have been obtained from systems where the agents are arranged in regular lattices, it is straightforward to apply our model to other topologies. In the Supplemental Material section A [25] we have included figures showing how the model behaves

on both random regular graphs and Erdős-Rényi (binomial) random graphs. The main results are very similar to those obtained on regular graphs, but there are a few differences worth mentioning. One difference is that without the low-dimensional spatial grid structure, the model has an even stronger tendency to end with almost all agents applying the same strategy. This results in some expanded bistable parameter regions around the transitions between the different phases. In these regimes, the systems will end up being dominated by one single strategy, but it is unpredictable from which strategy class. Another difference is caused by the variation of connectivities in the Erdős-Rényi graphs. Agents with more neighbors can potentially end up with higher payoffs. Since the model has a nonlinear tendency to select "the richest" neighbors, the dynamics tend to being dominated by the local structures around highly connected agents. This can lead to scattered strategy distributions when the selection pressure is high, $w \gtrsim 0.1$.

IV. DISCUSSION

In this paper we have studied a spatial prisoner's dilemma model in which agents marked with meaningless labels imitate the strategies of their neighbors—preferably the richest ones. If all agents are identical, i.e. in the special case with only $L = 1$ label, it is a well-known result discovered by Ohtsuki *et al.* [23], that cooperation is the evolutionarily stable strategy when the cooperation benefit is greater than the average connectivity, $b > k$. Our simulations confirm that this simple rule also applies to systems with a multitude of different labels, as long as the selection pressure is sufficiently low. We have also confirmed that the dynamics can lead to macroscopic states of persistent hierarchical discrimination, which was previously only demonstrated in the case of binary

labels ($L = 2$) [21]. This is an example of spontaneous symmetry breaking, since the model definition does not specify any differences between the labels. We have investigated systems with up to $L = 7$ different labels, and we found that the transition points between unconditional and discriminating strategies are essentially independent of the number of labels. This quantitative robustness is interesting when considering that the two unconditional strategies (cooperate or defect against all) constitute a rapidly decreasing fraction of the number of possible strategies, which increases exponentially with the number of labels.

While the main results known from the cases with $L = 1$ and $L = 2$ are essentially unchanged in systems with $L > 2$, our investigations have also revealed some new phenomena. One finding is that the parameter region dominated by discriminating strategies is subdivided into distinct phases dominated by different strategy classes. While one cannot know from the onset of a simulation which strategies are going to dominate in the system, it is possible to make reliable predictions about how many of the labels these strategies treat with cooperation and defection. This is similar to two discriminating phases "cooperate with the majority" and "cooperate with the minority" which were detected in the binary-label model.

Another fascinating phenomenon which can only be observed when $L > 2$ is that neighbors are more likely to have similar strategies. Since strategies are copied from neighbor to neighbor, it is no surprise that the model forms patches of agents agreeing on a single strategy. There are, however, no mechanisms in our model which explicitly favor neighbors with similar strategies. It is therefore surprising to observe that it is more likely for neighbors to have more similar strategies, when we measure similarity between two strategies as the number of label toward which they agree on what action to take. We do not yet understand this observation and finding an explanation will require further research.

Our system shows qualitatively similar results when the agents are arranged on random regular graphs, or even Erdős-Rényi graphs, rather than square lattices (as demonstrated in the Supplemental Material [25]). However, it is worth pointing out that it is necessary to enforce some kind of population structure for the model to express nontrivial dynamics. In a well-mixed system—i.e., on a fully connected graphs—the defect-all strategy would always be evolutionarily stable. The purpose of the paper is not to demonstrate a new mechanism to overcome the tragedy of the commons but rather to investigate the process by which discrimination can reduce cooperation asymmetrically in the presence of arbitrary labels.

The model described in this paper is not the only one in the literature of evolutionary game theory which is designed to investigate the combined effect of spatial structure and distinguishable agents. One example is the work by García *et al.* [15] who demonstrated that introducing tags may reduce the amount of cooperation in a structured population through

a mechanism they have called "the evil green-beard effect." The cooperation-reducing effect of tags through negative discrimination is closely related to results presented in this paper, but the dynamics leading to this effect is quite different. Our model tends to end up in a stationary state, whereas theirs exhibits cyclic behavior which is characteristic for tag-based cooperation [11–16,20]. One of the main differences is that the models of tag-based cooperation treat the tags as a part of the variable state subject to the evolution dynamics. Our labels, on the other hand, are immutable properties of the agents, determined at the beginning of the simulation. This can be seen as an approximation of a system in which behaviors adapt on a much faster timescale than physical appearances as, for example, if the dynamics are thought of as a type of learning via social imitation, while the labels represent genetically determined physical traits. This is reminiscent of the close connection between multiagent reinforcement learning and evolutionary dynamics [31,32].

Another important assumption in our model is that agents choose which of their neighbors to imitate solely based on their fitness. In particular, it does not consider the labels—neither its own nor those of the neighbors—or how they interact with the imitated strategy. This can be interpreted as if the agents are not aware of their own labels. It is difficult to imagine that humans should carry easily observable markers without being aware of them themselves. Our softer interpretation is that the strategies represent subconscious biases exempted from rational reasoning. Experimental evidence suggesting that humans do exhibit ingroup devaluation (or outgroup favoritism) [4,5,7] supports this viewpoint and finds a possible representation in the model.

When our model dynamics are interpreted as agents attempting to imitate the behaviors of their most successful neighbors, then the model parameter w can be interpreted as a kind of inverse temperature controlling noise level in the system. When the system is very "hot" (low selection pressure $w \to 0$) no strategies are preferred over any other, so the system will end up in a mixed state where all strategies are equally present. When the system is very cold (high selection pressure), we also observe a high coexistence of strategies. In the intermediate regime, however, we find that one strategy is likely to dominate the entire system. A similar nonmonotonic relation between noise and the degree of coexistence has been observed in spatial Lotka-Volterra systems with continuous noise [33,34].

In conclusion, the model we investigated in this paper demonstrates that imitating successful behaviors may lead to the emergence of persistent hierarchical discrimination in a population where agents are marked with observable, but otherwise meaningless, labels. We found this to be a remarkably robust phenomenon with respect to the number of labels. A central mechanism of the emergence of hierarchical social structures in the model is spontaneous symmetry breaking, transforming initial randomness into persistent fates.

[1] M. B. Brewer, The psychology of prejudice: Ingroup love and outgroup hate? J. Soc. Issues **55**, 429 (1999).

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

[3] C. L. Ridgeway, E. H. Boyle, K. J. Kuipers, and D. T. Robinson, How do status beliefs develop? the role of resources and inter-actional experience, Am. Sociol. Rev. **63**, 331 (1998).

[4] L. A. Rudman, J. Feinberg, and K. Fairchild, Minority members' implicit attitudes: Automatic ingroup bias as a function of group status, Soc. Cogn. **20**, 294 (2002).

[5] J. T. Jost, B. W. Pelham, and M. R. Carvallo, Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups, J. Exp. Soc. Psychol. **38**, 586 (2002).

[6] J. T. Jost, M. R. Banaji, and B. A. Nosek, A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo, Pol. Psychol. **25**, 881 (2004).

[7] D. S. March and R. Graham, Exploring implicit ingroup and outgroup bias toward hispanics, Group Process. Intergroup Relat. **18**, 89 (2015).

[8] A. Proestakis and P. Brañas-Garza, Self-identified obese people request less money: A field experiment, Front. Psychol. **7**, 1454 (2016).

[9] J. M. Smith, The theory of games and the evolution of animal conflicts, J. Theor. Biol. **47**, 209 (1974).

[10] M. A. Nowak, Five rules for the evolution of cooperation, Science **314**, 1560 (2006).

[11] R. L. Riolo, M. D. Cohen, and R. Axelrod, Evolution of cooperation without reciprocity, Nature **414**, 441 (2001).

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

[13] A. Traulsen and M. A. Nowak, Chromodynamics of cooperation in finite populations, PLOS One **2**, e270 (2007).

[14] E. Cohen, The evolution of tag-based cooperation in humans, Curr. Anthropol. **53**, 588 (2012).

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

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

[17] H. Zhang, Moderate tolerance promotes tag-mediated cooperation in spatial prisoner's dilemma game, Physica A (Amsterdam) **424**, 52 (2015).

[18] T. Wu, L. Wang, and F. Fu, Coevolutionary dynamics of phenotypic diversity and contingent cooperation, PLoS Comput. Biol. **13**, e1005363 (2017).

[19] T. Wu, F. Fu, and L. Wang, Phenotype affinity mediated interactions can facilitate the evolution of cooperation, J. Theor. Biol. **462**, 361 (2019).

[20] F. Fu, C. E. Tarnita, N. A. Christakis, L. Wang, D. G. Rand, and M. A. Nowak, Evolution of in-group favoritism, Sci. Rep. **2**, 460 (2012).

[21] G. G. Jensen and S. Bornholdt, Imitating the winner leads to discrimination in spatial prisoner's dilemma model, Sci. Rep. **9**, 3776 (2019).

[22] M. A. Nowak and R. M. May, Evolutionary games and spatial chaos, Nature **359**, 826 (1992).

[23] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks, Nature **441**, 502 (2006).

[24] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno, Evolutionary dynamics of group interactions on structured populations: A review, J. Roy. Soc. Interface **10**, 20120997 (2013).

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.100.062302 for simulation results using random graphs, a description of the network topology of lattices with different connectivities, and more examples concerning the likelihood of neighbors having similar strategies.

[26] P. Clifford and A. Sudbury, A model for spatial conflict, Biometrika **60**, 581 (1973).

[27] J. T. Cox and D. Griffeath, Diffusive clustering in the two dimensional voter model, Ann. Probab. **14**, 347 (1986).

[28] J.-M. Drouffe and C. Godrèche, Phase ordering and persistence in a class of stochastic processes interpolating between the ising and voter models, J. Phys. A **32**, 249 (1999).

[29] L. Dall'Asta and C. Castellano, Effective surface-tension in the noise-reduced voter model, Europhys. Lett. **77**, 60005 (2007).

[30] D. Volovik and S. Redner, Dynamics of confident voting, J. Stat. Mech.: Theory Exp. (2012) P04003.

[31] T. Börgers and R. Sarin, Learning through reinforcement and replicator dynamics, J. Econ. Theory **77**, 1 (1997).

[32] R. Nicole, P. Sollich, and T. Galla, Stochastic evolution in populations of ideas, Sci. Rep. **7**, 40580 (2017).

[33] A. Fiasconaro, D. Valenti, and B. Spagnolo, Nonmonotonic behavior of spatiotemporal pattern formation in a noisy lotka-volterra system, Acta Phys. Pol. B **35**, 1491 (2004).

[34] D. Valenti, A. Fiasconaro, and B. Spagnolo, Pattern formation and spatial correlation induced by the noise in two competing species, Acta Phys. Pol. B **35**, 1481 (2004).