Slow emergence of cooperation for win-stay lose-shift on trees

Elchanan Mossel · Sébastien Roch

Abstract We consider a group of agents on a graph who repeatedly play the prisoner's dilemma game against their neighbors. The players adapt their actions to the past behavior of their opponents by applying the win-stay lose-shift strategy. On a finite connected graph, it is easy to see that the system learns to cooperate by converging to the all-cooperate state in a finite time. We analyze the rate of convergence in terms of the size and structure of the graph. Dyer et al. (2002) showed that the system converges rapidly on the cycle, but that it takes a time exponential in the size of the graph to converge to cooperation on the complete graph. We show that the emergence of cooperation is exponentially slow in some expander graphs. More surprisingly, we show that it is also exponentially slow in bounded-degree trees, where many other dynamics are known to converge rapidly.

Keywords Games on graphs · Learning · Prisoner's dilemma game · Win-Stay Lose-Shift · Oriented percolation · Emergence of cooperation

1 Introduction

We consider a group of agents arranged on the nodes of a graph who repeatedly play the prisoner's dilemma game against their immediate neighbors. The players adapt their actions to the past behavior of their opponents by applying the so-called win-stay lose-shift strategy (Nowak & Sigmund, 1993) which, as the name suggests, consists in changing strategy whenever the payoff is deemed unsatisfactory. This model has been studied in the artificial intelligence literature (Kittock, 1995) as a simple example of "co-learning" (Shoham &

E. Mossel (✉) · S. Roch
Department of Statistics, University of California, Berkeley, Berkeley, CA 94720-3860
e-mail: mossel@stat.berkeley.edu

S. Roch
e-mail: sroch@stat.berkeley.edu

Tennenholtz, 1993, 1997). On a finite connected graph, it turns out that the system converges to the all-cooperate state—the globally optimal state—in finite time. In this respect, this instance of the iterated prisoner's dilemma (IPD) game on a graph provides an interesting example of a system learning to behave optimally by a mechanism that involves each agent applying independently a simple strategy—or rule of thumb—which takes into account only the latest actions of its immediate neighbors. For related work, see (Fudenberg & Levine, 1998) and references therein. See also (Axelrod, 1984) for the evolutionary perspective.

In order to understand how persistent this "emergence of cooperation" phenomenon is, it is crucial to analyze the rate of convergence to the all-cooperate state. Where the convergence is rapid, one would expect to observe the optimal, cooperation state in a practical system based on similar dynamics. On the other hand, where the convergence is slow, one would rather expect that such a system would stagnate in a suboptimal, metastable state where a nonnegligible fraction of agents defect. Rates of convergence for IPD were studied in Kittock (1995) and Dyer et al. (2002) where the structure of the graph was shown to be a determining factor.

In this paper, we show that IPD exhibits an exponentially slow convergence to cooperation on expander graphs and bounded-degree trees. Our result for bounded-degree trees is somewhat surprising. In particular, it should be compared to the behavior of global reversible dynamics on trees (Berger et al., 2005) where the convergence is always rapid. Note however that this slow convergence on trees is not unprecedented. Notably, the contact process, a common model of infection, is slow to converge on trees when the infection rate is large. See e.g. (Liggett, 1999) and references therein. In fact, our proof suggests that IPD behaves very much like the contact process. Nevertheless, the analysis of non-reversible particle systems has been an open challenge in the last two decades and we hope that the results obtained here can shed some more light on how such systems can be tackled.

The proof of slow convergence we give here combines several ideas. The main idea is to look at the process at the right space-time scaling. This approach, commonly used in probability (e.g. in the analysis of interacting particle systems (Liggett, 1985)), allows us to analyze the rough behavior of IPD—defection survives for long periods of time in zones that are densely populated by defectors. The main technical difficulty is to control the dependencies between different regions and different times. Then the process is compared to a directed percolation process (where the directed axis corresponds to the time axis in the original process). Using contour arguments we show that the directed percolation process survives for an exponential time. See (Durrett, 1984) for background on directed percolation.

1.1 Definitions and previous work

Recall that the prisoner's dilemma game (PD) is a bimatrix game with the following payoff matrix for the row player (and similarly for the column player):

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix}$$

where $T > R > P > S$ and $2R > T + S$. The first row (column) corresponds to the *cooperate* action and the second row (column) corresponds to the *defect* action. The global—or Pareto—optimum is for both agents to cooperate. However, for any given action of the col-

umn player, it is always in the row player's advantage to defect (and similarly for the column player).

For an agent playing PD, a simple way to adapt to her opponent's behavior is the so-called Win-Stay Lose-Shift strategy (WSLS) (Nowak & Sigmund, 1993), also known as the Pavlov strategy (Kittock, 1995; Shoham & Tennenholtz, 1997). This works as follows. Every time the game is played, if the agent's payoff is one of the two smaller payoffs, i.e. P or S, then she switches her action in anticipation for the next round of play, otherwise she keeps the same action.

We now consider a repeated graphical version of PD which we will refer to as IPD. Let $G = (V, E)$ be a finite graph with $n = |V|$. Each node, v, is an agent to which we associate an action $A_t(v) \in \{C, D\}$ at time $t \in \mathbb{R}_+$. Here C stands for *cooperation* while D stands for *defection*. The initial state is $A_0(v) = D$ for all $v \in V$. The agents repeatedly play PD against their immediate neighbors in the graph through the following mechanism. Each edge $e \in E$ has an exponential clock, i.e. we associate to each edge an independent Poisson process $\{T_i(e)\}_{i \geq 1}$ where all inter-arrival times $T_{i+1}(e) - T_i(e)$ are independent Exp(1) (with the convention $T_0 = 0$). Every time a clock rings, say at edge $e = (u, v)$, the endpoint agents u and v play one round of PD using their respective actions $A_t(u)$ and $A_t(v)$, assuming the clock rings at time t. Then the two agents update their state using WSLS. In other words, if a clock rings on edge $e = (u, v)$ at time t, we witness the following transition for $(A_t(u), A_t(v))$

$$(C, C) \rightarrow (C, C)$$
$$(C, D) \rightarrow (D, D)$$
$$(D, C) \rightarrow (D, D)$$
$$(D, D) \rightarrow (C, C).$$

Equivalently, we could consider a discrete-time process where at each time step, one edge is picked uniformly at random among all edges and is updated as above. The correspondence between the two definitions is well-known (see e.g. (Aldous & Fill, 2006)). In particular, note that in the continuous-time case, no two clocks can ring simultaneously because the interarrival times have a continuous density. Moreover, by the memoryless property of the exponential, any time a clock rings, the next clock to ring is uniform over all edges. The discrete-time version is used in Dyer et al. (2002) and in Section 3 of the present paper. But, as will become clear below, it is easier to consider the continuous-time version of IPD in Section 2.

The update mechanism described above defines a stochastic process for the state of the system $\mathbf{A}_t = (A_t(v))_{v \in V}$ with initial state the all-defect state, $\mathbf{A}_0 = \mathbf{D} \equiv (D, \ldots, D)$. It is clear that, given the above allowed transitions, the system has a unique fixed point, the all-cooperate state $\mathbf{C} \equiv (C, \ldots, C)$. In particular, if G is a finite connected graph with $n \geq 2$, we have a.s. $\mathbf{A}_t \rightarrow \mathbf{C}$ as $t \rightarrow +\infty$. The question of interest is: how long does it take to reach $\mathbf{C}$ on a given graph. It was shown by Dyer et al. (2002)—and previously conjectured in Kittock (1995)—that the time to the emergence of cooperation depends crucially on the structure of the graph. Let $T_{\mathbf{C}}$ be the stopping time at which $\mathbf{A}_t$ reaches $\mathbf{C}$ for the first time. Below, *with high probability* (w.h.p.) means with probability $1 - 1/\text{poly}(n)$ where poly(n) increases polynomially with n. In Dyer et al. (2002), the following two results are proved.

Theorem 1. (Dyer et al., 2002) *Let G be a cycle on n vertices. Then w.h.p. $T_{\mathbf{C}} = O(n \log n)$.*

Theorem 2. (Dyer et al., 2002) *Let G be the complete graph on n vertices. Then w.h.p.* $T_{\mathbf{C}} = \Omega((1.1)^n)$.

1.2 Our results

Given the previous theorems, it is natural to conjecture that the time to the emergence of cooperation is governed by the connectivity of the graph: a high connectivity, as in the complete graph, leads to slow convergence, while a low connectivity, as in the cycle, leads to fast convergence. Surprisingly, we refute this intuition with our main result.

Theorem 3. *There is a constant d so that for all n there is a d-regular tree with n vertices for which w.h.p.* $T_{\mathbf{C}} = \Omega(\rho^n)$ *for some* $\rho > 1$ *depending only on d.*

To prove this result, we study IPD on "linear trees." The main technical ingredient is a coupling with oriented percolation. The proof of this theorem is given in Section 2.

Although the connectivity conjecture turns out to be wrong in general, the following theorem, an extension of the complete graph result of Dyer et al. (2002), shows that the intuition is partly correct in one direction. More precisely, Theorem 4 below guarantees that IPD is slow to converge on any graph with high enough connectivity while Theorem 3 above shows that IPD can be slow even if connectivity is very low. Let G be a graph with n vertices. Let α, β be two increasing functions of n such that for all n, $0 < \alpha(n) < \beta(n) < n$. Define the (α, β)-expansion constant $\rho_{\alpha,\beta}(G)$ of G as

$$\rho_{\alpha,\beta}(G) = \min \left\{ \frac{|E(U, U^c)|}{\text{vol}(U)} \ : \ U \subseteq V, \ \alpha(n) \le |U| \le \beta(n) \right\},$$

where $E(U, U^c)$ is the set of edges between U and U^c, $\text{vol}(U)$ is the sum of the degrees of the nodes in U, and $|X|$ is the cardinality of X.

Theorem 4. *Let* $\varepsilon > 0$. *Let* α, β *be two increasing functions of n such that for all n,* $0 < \alpha(n) < \beta(n) < n$. *Let G be a graph with n vertices such that* $\rho_{\alpha,\beta}(G) > 1/2 + \varepsilon$. *Then there is a constant* $a > 1$ *(depending only on* ε*) such that w.h.p.* $T_{\mathbf{C}} = \Omega(a^{\beta(n)-\alpha(n)})$ *(for n large enough). In particular, if* α, β *are linear in n, the emergence of cooperation is exponentially slow.*

This follows from a martingale argument similar to that used in Dyer et al. (2002) which is detailed in Section 3. Note that in Theorem 4, in order to obtain slow convergence, it suffices to have large expansion for relatively small sets. In particular, the theorem applies to expander graphs such as random regular graphs (Kahale, 1995; Friedman, Kahn, & Szemeredi, 1989).

2 Win-stay lose-shift on trees

In this section, we analyze IPD on caterpillar trees of degree d. We define an *(n,d)-caterpillar*, denoted $\mathbb{S}_d^n$, to be a tree with the following property: the subtree induced by the internal nodes is a path containing n nodes all of which have degree d. See Fig. 1. Our main result, Theorem 3, is that cooperation is slow to emerge on caterpillars. The proof of Theorem 3 follows from a series of stochastic domination arguments. We now briefly outline the main steps of the proof.
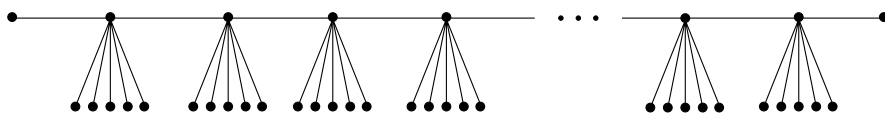
Fig. 1 Caterpillar of degree 7

1. **Star Dynamics via Biased Random Walk.** The first step is to analyze the behavior of a single star. The main point here is that it takes the star with d leaves an exponential number of steps (in d) to move from the all-defect state to the all-cooperate state. This is proved by comparing the process to a biased random walk. This comparison also shows that a star can go from a few defectors to linearly many in poly(d) time with constant probability, and that a small linear fraction of defectors grows with high probability within poly(d) steps. Moreover, these claims can be established even if one allows two of the nodes of the stars to have arbitrary values.

2. **Space-Time Scaling.** We think of a star as defecting if at least $d/4$ of its leaves defect. Then, we consider triplets of adjacent stars and say that a triplet is defecting if at least one of its extremal stars is defecting. (We actually work with triplets of stars rather than pairs to help control dependencies.) We scale time by looking at the process every poly(d) steps. The random walk argument of the previous point allows to show that defecting stars have a high probability—at least $(1 - \exp(-\Omega(d)))$—of remaining defectors after the poly(d) time window. Moreover, a defecting star has a $1/\text{poly}(d)$ probability of "infecting" neighboring stars during that time. By iterating these observations poly(d) times—yet another time scaling—we show that a defecting triplet has a probability $1 - \exp(-\Omega(d))$ of "infecting" a neighboring triplet. (Neighboring triplets are actually intersecting.)

3. **Percolation.** We may now look at the space-time diagram of defecting triplets and show that it dominates a directed percolation with probability $1 - \exp(\Omega(-d))$ for edges to be open. The time axis of the original process corresponds to the direction of propagation in the percolation process. Finally, a contour argument allows to conclude that this percolation survives for a time which is exponential in n, thus proving that the convergence time of IPD on the caterpillar is itself exponential in n.

2.1 Star

Let $G = \mathbb{S}_d^n$. This graph is made of n copies of $\mathbb{S}_d^1$ (i.e. stars of degree d). Let G' be any star in G. Denote the root 0 and the leaves $1, 2, \ldots, d$. A crucial property of stars is that cooperation is slow to emerge on them. This follows from our next result. We single out nodes 1 and 2, which are defined to be the two nodes that G' shares with its neighboring stars. (In the case of extremal stars, we just pick an arbitrary node in addition to the node shared with the next star.) We call 1 and 2 the external vertices. We use the following notation: $a \wedge b = \min\{a, b\}$.

Lemma 1 (Dynamics on Stars). *Consider the IPD chain $\{\mathbf{A}_t\}_{t \geq 0}$ on $G = \mathbb{S}_d^n$ with $d > 15$. Let G' be an arbitrary star in G with nodes denoted $0, \ldots, d$ (0 being the root, and 1 and 2 being the external vertices). Let M' be a positive integer and g_0, g_1, g_2 be three increasing functions of d with $g_2(d) = d/3 - 2$ and g_0, g_1 satisfying $1 < g_0(d) < g_1(d) < g_2(d)$ for all d. Let the initial configuration be as follows. On G', nodes 3 through $d - g_1$ are C and nodes $d - g_1 + 1$ through d are D. On all other nodes, including the root and external vertices of*

G', *the initial state is arbitrary. Define*

$$N_D = |\{i \in \{3, \dots, d\} \; : \; A(i) = D\}|.$$

Let T_g be the first time $N_D = g$. Let $\Delta_2 = g_2 - g_1$, $\Delta_1 = g_1 - g_0$, $\rho = \sqrt{9/8}$, and $\mu = g_0 M'$. Then, we have

$$\mathbb{P}[T_{g_2} \geq (T_{g_0} \wedge M')] \leq 2^{-\Delta_1} + \rho^{-\sqrt{\mu}/2}(\sqrt{2})^{\Delta_2} + 2^{-\mu/2}. \tag{1}$$

Moreover, this bound applies simultaneously on all stars independently from each other (possibly with different choices of g's).

Proof: For this argument, we restrict ourselves to what happens on G' and do not refer to any event involving the rest of G. We call a leaf edge with leaf state D a D-edge, and similarly for C. The behavior of N_D depends on the state at the root of G'. When $A(0) = C$, nothing happens until a D-edge is picked at which time $A(0)$ becomes D itself. On the other hand, when $A(0) = D$, either a C-edge is chosen in which case N_D may go up by 1 (or stay the same if 1 or 2 is picked), or a D-edge is chosen in which case N_D may go down by 1 (or stay the same if 1 or 2 is picked) and $A(0)$ becomes C. Ignore the updates where nothing changes, i.e. when an edge (C, C) is chosen. In any configuration satisfying $N_D \geq g_0$, there are at least g_0 edges whose updates change the configuration. Let Q the number of such updates in time M'. Then it follows that Q is larger than a Poisson with mean $\mu = g_0 M'$. From the moment generating function of the Poisson distribution (see e.g. (Durrett, 1996)), we have the following

$$\mathbb{P}[Q \leq \sqrt{\mu}] = \mathbb{P}[e^{\mu - Q} \geq e^{\mu - \sqrt{\mu}}] \leq \frac{\mathbb{E}[e^{\mu - Q}]}{e^{\mu - \sqrt{\mu}}} \leq \frac{e^{\mu} e^{\mu(e^{-1} - 1)}}{e^{\mu - \sqrt{\mu}}} \leq 2^{-\mu/2},$$

for μ large enough. Assume the event $\{Q \geq \sqrt{\mu}\}$ holds. Also, note that at most one out of 2 steps have $A(0) = C$. (Remember that we ignore (C, C) updates.) Ignore the times with $A(0) = C$ as well, what remains is an asymmetric random walk (or rather a birth-and-death chain) which does at least $\sqrt{\mu}/2$ steps before time M'. To bound the probability that N_D goes up or down, we use the fact that the chain starts with g_1 D's and is stopped when it reaches either g_0 or g_2 D's. By assumption, the probability that N_D goes up when $A(0) = D$ is at least $(d - 2 - g_2)/d$. Consider the walk $\{S_k\}_{k \geq 0}$ on $\mathbb{N}$ started at $S_0 = g_1$ which goes up with probability $p = (d - 2 - g_2)/d = 2/3$ and goes down with probability $1 - p = 1/3$. Let T'_g be the time at which S_k reaches g. For convenience, we assume that the process $\{S_k\}_{k \geq 0}$ is defined on all of $\mathbb{Z}$ (even though outside the interval $[g_0, g_2]$ the bounds used are not valid). Then,

$$\mathbb{P}[T_{g_2} \geq (T_{g_0} \wedge M') \mid Q \geq \sqrt{\mu}] \leq \mathbb{P}[T'_{g_2} \geq (T'_{g_0} \wedge \sqrt{\mu}/2)]$$
$$\leq \mathbb{P}[T'_{g_2} \geq T'_{g_0}] + \mathbb{P}[T'_{g_2} \geq \sqrt{\mu}/2].$$

By standard martingale results (see e.g. (Durrett, 1996, Example 4.7.1)), we have

$$\mathbb{P}[T'_{g_2} \geq T'_{g_0}] = \frac{\phi(\Delta_2) - \phi(0)}{\phi(\Delta_2) - \phi(-\Delta_1)},$$

where

$$\phi(s) = \left(\frac{1-p}{p} \right)^s = 2^{-s}.$$

So,

$$\mathbb{P}[T'_{g_2} \geq T'_{g_0}] = \frac{1 - 2^{-\Delta_2}}{2^{\Delta_1} - 2^{-\Delta_2}} \leq 2^{-\Delta_1}.$$

We also have

$$\mathbb{E}[\rho^{T'_{g_2}}] = \left(\frac{1 - \sqrt{1 - 4p(1-p)\rho^2}}{2(1-p)\rho} \right)^{\Delta_2}.$$

The choice $\rho = \sqrt{9/8}$ gives

$$\mathbb{E}[\rho^{T'_{g_2}}] = (\sqrt{2})^{\Delta_2}.$$

By Markov's inequality,

$$\mathbb{P}[T'_{g_2} \geq \sqrt{\mu}/2] = \mathbb{P}[\rho^{T'_{g_2}} \geq \rho^{\sqrt{\mu}/2}] \leq \rho^{-\sqrt{\mu}/2} \left(\sqrt{2} \right)^{\Delta_2}.$$

Finally, putting everything together, we get (1).

The independence of the bound at each star in G comes from the fact that we use only events involving leaf edges of G'. □

The following corollary corresponds to the case where a star has initially only a few D's. The result below implies that after $M' = \text{poly}(d)$ steps the star has $O(d)$ D's with positive probability.

Corollary 1 (Defection Spreads on Stars). *In the setup of Lemma* 1, *let* $g_0(d) = 2$, $g_1(d) = 3$ *and* $g_2(d) = d/3 - 2$. *Then, for* $M' = \omega(d^2)$ *and* d *(constant) large enough, we have*

$$\mathbb{P}[T_{g_2} \geq (T_{g_0} \wedge M')] \leq \frac{2}{3}.$$

The following corollary implies that a star with $O(d)$ D's still has $O(d)$ D's after $\text{poly}(d)$ steps, with high probability.

Corollary 2 (Defection Survives on Stars). *In the setup of Lemma* 1, *let* $M' \to +\infty$, $g_2(d) = d/3 - 2$, $g_1(d) = d/3 - 3$, *and* $g_0(d) = d/4 - 3$. *Then,*

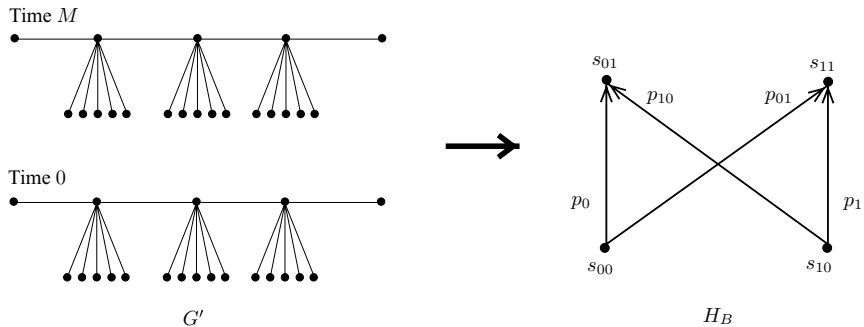$$\mathbb{P}[T_{g_2} \geq T_{g_0}] \leq 2^{-d/12}.$$

Fig. 2 Reduction to percolation. The analysis of IPD on G' (left) uses a percolation model on H_B (right). The states $s_{00}, s_{10}, s_{01}, s_{11}$ are in $\{0, 1\}$, and the values p_0, p_1, p_{10}, p_{01} indicate the probability that the corresponding edge is open. Each of the top states is 1 if it is the endpoint of an open edge with bottom state 1, and is 0 otherwise

The following corollary implies that a star with $d/4$ D's reaches $d/3$ D's after poly(d) steps, with high probability.

Corollary 3 (Defection Boosting on Stars). *Let τ be a positive integer, not depending on d. In the setup of Lemma* 1, *let $g_2(d) = d/3 - 2$, $g_1(d) = d/4 - 2 - \tau$, and $g_0(d) = d/5 - 2 - \tau$. Then, for $M' = \omega(d^2)$ and d large enough, we have*

$$\mathbb{P}[T_{g_2} \geq (T_{g_0} \wedge M')] \leq 3\, 2^{-d/20} \leq 2^{-d/21}.$$

2.2 Star triplets

The next step in the proof of Theorem 3 is to make the connection between IPD and oriented percolation. Here we show how a triplet of stars dominates the building block of a percolation lattice. We use the following oriented percolation. Consider four adjacent vertices of the regular lattice $\mathbb{Z}^2$, say $v_{00} = (0, 0)$, $v_{01} = (0, 1)$, $v_{10} = (1, 0)$ and $v_{11} = (1, 1)$. Assume the nodes are connected by four directed edges: $e_0 = (v_{00}, v_{01})$, $e_1 = (v_{10}, v_{11})$, $e_{01} = (v_{00}, v_{11})$, and $e_{10} = (v_{10}, v_{01})$. See Fig. 2. Each edge is *open* with respective probability p_0, p_1, p_{01}, and p_{10}. The vertices have a state, denoted respectively $s_{00}, s_{01}, s_{10}, s_{11}$, which takes its value in $\{0, 1\}$. The state 1 "travels" along the open edges, i.e. if $e = (u, v)$ is an open edge and the state at u is 1 then the state at v is also 1. A vertex is in state 1 if and only if it is the terminal vertex of an open edge with initial vertex in state 1. We denote this four-node graph H_B.

Now consider any triplet of adjacent stars inside $G = \mathbb{S}_d^n$. Denote the stars S_j, $j = 1, 2, 3$, with corresponding edges $\{e_i^{(j)}\}_{j=1}^d$ and vertices $\{v_i^{(j)}\}_{i=0}^d$, with the label 0 corresponding to the root. We have the correspondence $e_2^{(1)} = e_1^{(2)}$ and $e_2^{(2)} = e_1^{(3)}$. We denote this subgraph— which is a copy of $\mathbb{S}_d^3$—G'. We are interested in the number of D's on each star, excluding nodes 0, 1, and 2 of each star, which we denote $\mathbf{N}_t = (N_t^{(1)}, N_t^{(2)}, N_t^{(3)})$.

The detailed behavior of $\mathbf{N}_t$ is rather intricate. We simplify the process by projecting it to a smaller space. Let

$$\sigma_d[N] = \begin{cases} 1, & \text{if } N > d/4 - 2, \\ 0, & \text{if otherwise.} \end{cases}$$

Consider the random vector

$$\tilde{\mathbf{s}} = (\tilde{s}_{00}, \tilde{s}_{01}, \tilde{s}_{10}, \tilde{s}_{11}) = \left(\sigma_d\big[N_0^{(1)}\big], \sigma_d\big[N_M^{(1)}\big], \sigma_d\big[N_0^{(3)}\big], \sigma_d\big[N_M^{(3)}\big]\right),$$

for some $M > 0$. The following lemma shows that for an appropriate choice of M, p_0, p_1, p_{01}, and p_{10}, the vector $\tilde{\mathbf{s}}$ stochastically dominates

$$\mathbf{s} = (s_{00}, s_{01}, s_{10}, s_{11}),$$

defined by the percolation above (with $s_{00} = \tilde{s}_{00}$ and $s_{10} = \tilde{s}_{10}$).

Lemma 2 (Connection to Percolation). *Consider the IPD chain $\{\mathbf{A}_t\}_{t \geq 0}$ on $G = \mathbb{S}_d^n$ with $d > 15$. Let G' be an arbitrary triplet of adjacent stars in G. Let $M = d^6$, $p_0 = p_1 = 1 - 2^{-d/30}$, and $p_{01} = p_{10} = d^{-10}$. Then, for any initial configuration and s_{00}, s_{10} such that $s_{00} = \tilde{s}_{00}$ and $s_{10} = \tilde{s}_{10}$, we have that $(\tilde{s}_{01}, \tilde{s}_{11})$ stochastically dominates (s_{01}, s_{11}) for d (constant) large enough. Moreover, the domination holds for any number of (edge-)nonintersecting triplets simultaneously independently from each other.*

Proof: The argument ignores any event outside G'. We consider three cases.

(1) *Case $\tilde{s}_{00} = \tilde{s}_{10} = 0$.* In that case, we have $s_{01} = s_{11} = 0$, which is of course dominated by $\tilde{s}_{01}, \tilde{s}_{11}$.

(2) *Case $\tilde{s}_{00} = \tilde{s}_{10} = 1$.* We use Corollaries 2 and 3, which we apply to stars 1 and 3 independently. Consider star 1. We first go through a "boosting" phase where we let $N^{(1)}$ drift from $d/4 - 2$ to $d/3 - 2$. Then we compute the probability that $N^{(1)}$ stays above $d/4 - 2$ for the remaining time.
Phase 1. For the boosting phase, we apply Corollary 3. The probability of remaining below $d/3 - 2$ is at most $2^{-d/21}$.
Phase 2. The time remaining after boosting is of course at most M. In time M, there is a Poisson number of steps, say Q', with mean dM (including the steps where nothing happens). From the moment generating function of the Poisson distribution (see e.g. (Durrett, 1996)), we have the following

$$\mathbb{P}[Q' \geq d^2 M^2] = \mathbb{P}\big[e^{Q'} \geq e^{d^2 M^2}\big] \leq \frac{\mathbb{E}[e^{Q'}]}{e^{d^2 M^2}} \leq \frac{e^{dM(e-1)}}{e^{d^2 M^2}} \leq 2^{-d^2 M^2/2}.$$

Assuming $d/3 - 2$ was reached and that there remains at most $d^2 M^2$ discrete steps, we get that there are at most $d^2 M^2$ crossings of the interval $[d/4 - 3, d/3 - 2]$ by the process $N^{(1)}$. By Corollary 2, every time $N^{(1)} = d/3 - 3$, there is a probability of at least $1 - 2^{-d/12}$ of coming back to $d/3 - 2$ before hitting $d/4 - 3$. The probability that any of $d^2 M^2$ attempts at crossing $[d/4 - 3, d/3 - 2]$ succeeds is at most at most $d^2 M^2 2^{-d/12}$ which implies

$$\mathbb{P}[\tilde{s}_{10} = 0] \leq d^2 M^2 2^{-d/12} + 2^{-d^2 M^2/2} + 2^{-d/21} \leq 2^{-d/22},$$

for d large enough. Stochastic domination of the oriented percolation follows directly.

(3) *Case $\tilde{s}_{00} = 1$, $\tilde{s}_{10} = 0$.* (The symmetric case is analyzed similarly.) We divide the time window in two phases. For the first phase, we compute the probability that defection "spreads" from star 1 to star 3. For the second phase, we compute the probability that stars 1 and 3 remain in or reach state 1 respectively.

Phase 1. It is easy to see that, in any initial configuration satisfying $\tilde{s}_{00} = 1$, $\tilde{s}_{10} = 0$, six steps (or less) suffice to reach a configuration with $N^{(3)} \geq 3$. The probability that the first six steps taken by IPD satisfy this property—call that event B—is at least $1/d^6$. Let Q'' be the number of steps until time $M/2$. Then,

$$\mathbb{P}[Q'' \leq 5] \leq 2^{-M/4},$$

by a calculation similar to that in Lemma 1.

Phase 2. We condition on $\{Q'' \geq 6\}$. Consider first star 1. Whether or not B is realized, at the beginning of Phase 2, we have $N^{(1)} \geq d/4 - 8$. We are back in the situation of Case 2), except that the time left is only at least $M/2$. By the same calculation, we obtain that the probability that $\tilde{s}_{10}$ is 0 is at most $2^{-d/22}$ for d large enough. Consider now star 3. Let Q''' be the number of discrete steps left on star 3. The time remaining is at least $M/2$. It follows from Corollary 1 that $N^{(3)}$ reaches $d/3 - 2$ before the end of the time window with probability at least $1/3$ for d large enough. Once $d/3 - 2$ is reached, we are back to Phase 2 of Case 2). It follows that on $\{Q'' \geq 6\}$ the probability that $\tilde{s}_{11} = 1$ is at least $d^{-6}/4$. Note that on $\{Q'' \geq 6\}$, the bounds on star 1 and 3 are independent. It is then easy to check that stochastic domination of the oriented percolation holds.

□

We further simplify the chain by stacking up the construction in the previous lemma and projecting once more to a smaller space. For this, we consider a different percolation model on $\mathbb{Z}^2$. See Fig. 3. Let H'_B be the directed graph made of three nodes $v'_{10} = (1, 0)$, $v'_{01} = (0, 1)$, $v'_{21} = (2, 1)$ with two edges $e'_1 = (v'_{10}, v'_{01})$, $e'_2 = (v'_{10}, v'_{21})$. The edges are open with
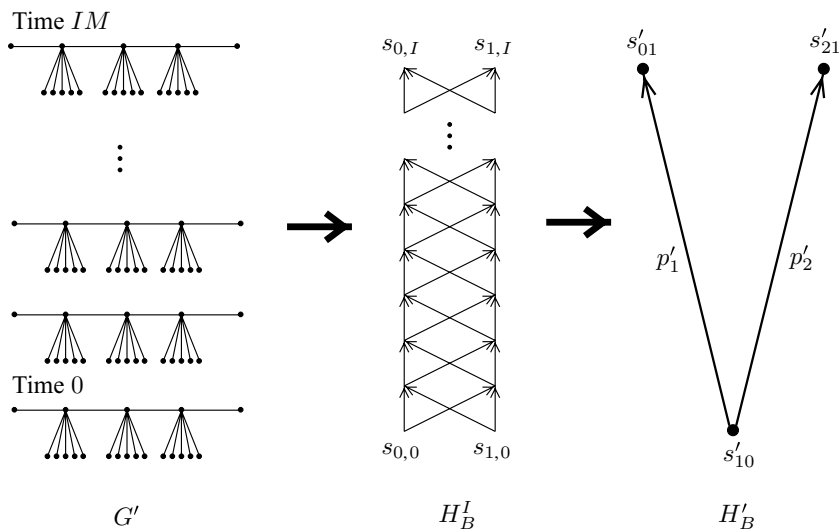

Fig. 3 Further reduction. The analysis of IPD on G' (left) at multiple times of IM is further reduced to a percolation on the simpler graph H'_B (right). Firstly, the reduction depicted in Fig. 2 is stacked up into a percolation on the tower H^I_B (middle). Then, the intermediate levels in H^I_B are combined to produce H'_B

probability p_1', p_2' respectively. The nodes have state s_{10}', s_{01}', s_{21}' respectively with value in $\{0, 1\}$. The percolation works as before with state 1 "traveling" along open edges.

Consider again IPD on an arbitrary triplet of stars G' of G. Redefine the vector $\tilde{\mathbf{s}}$ by taking instead

$$\tilde{\mathbf{s}} = (\tilde{s}_{00}, \tilde{s}_{01}, \tilde{s}_{10}, \tilde{s}_{11}) = \left(\sigma_d\big[N_0^{(1)}\big], \sigma_d\big[N_{IM}^{(1)}\big], \sigma_d\big[N_0^{(3)}\big], \sigma_d\big[N_{IM}^{(3)}\big] \right),$$

for some $I \in \mathbb{N}$ and M as in Lemma 2. We use the following notation: $a \vee b = \max\{a, b\}$.

Lemma 3 (Towers). *Consider the IPD chain $\{\mathbf{A}_t\}_{t \geq 0}$ on $G = \mathbb{S}_d^n$ with $d > 15$. Let G' be an arbitrary triplet of adjacent stars in G. Let $M = d^6$, $I = d^{100}$, and $p_1' = p_2' = 1 - 2^{-d/100}$. Then, for any initial configuration and s_{10}' such that $s_{10}' = \tilde{s}_{00} \vee \tilde{s}_{10}$, we have that $(\tilde{s}_{01}, \tilde{s}_{11})$ stochastically dominates (s_{01}', s_{21}') for d (constant) large enough. Moreover, the domination holds for any number of (edge-)nonintersecting triplets simultaneously independently from each other.*

Proof: The argument ignores any event outside G'. The proof works by stacking up I copies of H_B and applying Lemma 2. Consider again $\mathbb{Z}^2$. We define a I-*tower*, denoted H_B^I, to be the graph on nodes $\{v_{0,i} = (0, i), v_{1,i} = (1, i)\}_{i=0}^I$ where each set of four nodes of the form $\{v_{0,i}, v_{1,i}, v_{0,i+1}, v_{1,i+1}\}$ induces a copy of H_B with the same values of p_0, p_1, p_{10}, p_{01} as in Lemma 2. The node states are denoted $\{s_{0,i} = (0, i), s_{1,i} = (1, i)\}_{i=0}^I$. By applying repeatedly Lemma 2, we get that, if $(\tilde{s}_{00}, \tilde{s}_{10}) = (s_{0,0}, s_{1,0})$, then $(\tilde{s}_{01}, \tilde{s}_{11})$ stochastically dominates $(s_{0,I}, s_{1,I})$, so it suffices to show that the latter dominates (s_{10}', s_{21}').

The case $s_{10}' = 0$ is trivial. So assume $s_{10}' = 1$. Then, the subcase $\tilde{s}_{00} \wedge \tilde{s}_{10} = 1$ dominates the subcase $\tilde{s}_{00} \wedge \tilde{s}_{10} = 0$ so it suffices to consider the latter. Without loss of generality, let $\tilde{s}_{00} = 1$ and $\tilde{s}_{10} = 0$. The probability that at least one upwards edge in H_B^I is closed is at most

$$2I\left(2^{-d/30}\right) \leq 2^{-d/31},$$

for d large enough. The probability that no up-right edge is open is at most

$$\left(1 - \frac{1}{d^{10}}\right)^I \leq 2^{-d/31},$$
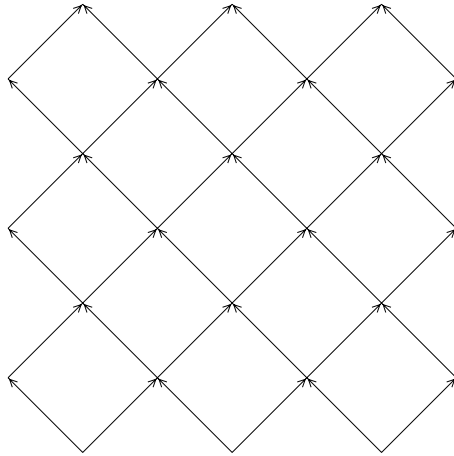
for d large enough. Therefore,

$$\mathbb{P}[s_{0,I} = s_{0,I} = 1] \geq 1 - 2^{-d/32},$$

for d large enough. But note that

$$\mathbb{P}[s_{01}' = s_{21}' = 0] = (2^{-d/100})^2 = 2^{-d/50} \geq 2^{-d/32}.$$

So we have domination. $\qquad\square$

Fig. 4 A section of the oriented percolation lattice used in the proof of Lemma 5

2.3 Oriented percolation

We conclude the proof of Theorem 3 by showing that the IPD chain at intervals of time IM dominates a standard percolation model and that in turn the latter model percolates at an exponential distance from its bottom nodes.

For convenience, assume n is of the form

$$n = 2n' + 1,$$

for some positive integer n'. (The reason for this choice will be clear below. See also Fig. 5.) Consider the following sublattice of $\mathbb{Z}^2$,

$$\mathcal{P} = \{(i, j) \in \mathbb{Z}^2 \; : \; 1 \leq i \leq n', \; 0 \leq j \leq T, \; i + j \text{ is even}\},$$

where T is a positive integer that will be fixed below. Consider the directed graph $G_{\mathcal{P}} = (V_{\mathcal{P}}, E_{\mathcal{P}})$ with node set $V_{\mathcal{P}} = \{v_{i,j}\}_{(i,j) \in \mathcal{P}}$ and edge set

$$E_{\mathcal{P}} = \{(v_{i,j}, v_{i+1,j+1}), (v_{i,j}, v_{i-1,j+1})\}_{(i,j) \in \mathcal{P}}.$$

See Fig. 4 for an illustration. Each edge has probability p' of being open where p' is set below. We consider the percolation process on $G_{\mathcal{P}}$ and denote the states $\mathbf{s}'_{\mathcal{P}} = \{s'_{i,j}\}_{(i,j) \in \mathcal{P}}$.

Let $\{\mathbf{A}_t\}_{t \geq 0}$ be the IPD chain on $\mathbb{S}^n_d$ and denote $N_t^{(i)}$ the number of D's on star i at time t, excluding the external nodes. We consider the following projection of $\{\mathbf{A}_t\}_{t \geq 0}$. Let

$$\mu(i, j) = 4(i - 1) + 1_{\{j \text{ is odd}\}},$$

and let $\tilde{\mathbf{s}} = \{\tilde{s}_{i,j}\}_{(i,j) \in \mathcal{P}}$ where

$$\tilde{s}_{i,j} = \sigma_d \left[N_{jIM}^{(\mu(i,j)-1)} \right] \vee \sigma_d \left[N_{jIM}^{(\mu(i,j)+1)} \right],$$

where I and M are as in Lemma 3. See Fig. 5. We show first that $\tilde{\mathbf{s}}$ dominates $\mathbf{s}'$.
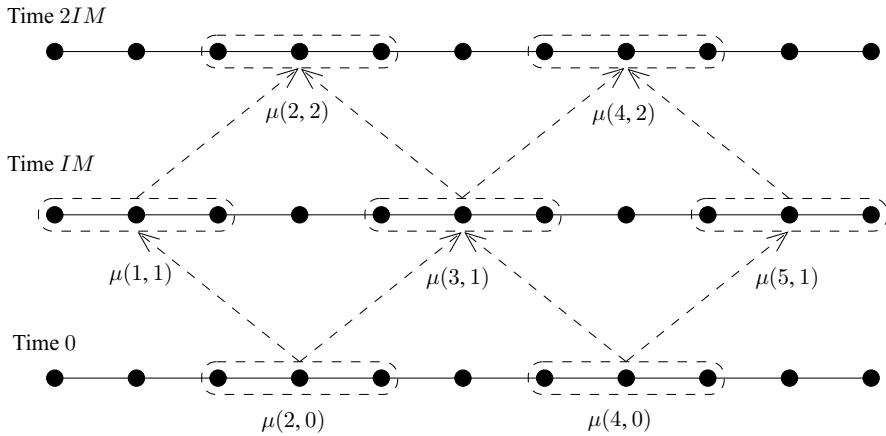
Fig. 5 Graph $\mathbb{S}_d^{11}$ (stars not shown) at times $0, IM, 2IM$. Here $n' = 5$. Circled triplets correspond to nodes of the percolation lattice of Fig. 4

Lemma 4 (Domination of Oriented Percolation). *Consider the IPD chain $\{\mathbf{A}_t\}_{t \geq 0}$ on $G = \mathbb{S}_d^n$ with $d > 15$. Let $M = d^6$, $I = d^{100}$, and $p' = 1 - 2^{-d/100}$. Let $\mathbf{A}_0 = \mathbf{D}$ (the all-D state) and let $s'_{i,0} = 1$ for all even i's. Then, we have that $\tilde{\mathbf{s}}$ stochastically dominates $\mathbf{s}'$ for d (constant) large enough.*

Proof: This actually follows immediately from Lemma 3. □

Finally, the next lemma concludes the proof of Theorem 3.

Lemma 5 (Crossing). *Let $\mathbf{s}'$ be defined as above with $p' = 1 - 2^{-d/100}$ and let $s'_{i,0} = 1$ for all even i's. Let $T = 2^{(d/2000)n}$. Assume that $n = 2n' + 1$ and that T is even. Then*

$$\mathbb{P}[s'_{i,T} = 0, \ \forall i \in \{2, 4, \ldots, n' - 1\}] \leq 2^{-(d/1000)n},$$

for d (constant) large enough.

Proof: We use a standard duality argument. For more details, see (Durrett, 1984). First we modify the percolation lattice $G_{\mathcal{P}}$, which we now call the primal lattice and still denote $G_{\mathcal{P}}$. To each edge, we add another edge, reversed, with associated probability of being open 0. We now define the dual lattice. Let

$$\mathcal{D} = \{(i, j) \in \mathbb{Z}^2 \ : \ 1 \leq i \leq n', \ 0 \leq j \leq T, \ i + j \text{ is odd}\}.$$

Consider the directed graph $G_{\mathcal{D}} = (V_{\mathcal{D}}, E_{\mathcal{D}})$ with node set $V_{\mathcal{D}} = \{v_{i,j}\}_{(i,j) \in \mathcal{D}}$ and edge set

$$E_{\mathcal{D}} = \{(v_{i,j}, v_{i+1,j+1}), (v_{i,j}, v_{i-1,j+1}), (v_{i,j}, v_{i-1,j-1}), (v_{i,j}, v_{i+1,j-1})\}_{(i,j) \in \mathcal{D}}.$$

Superimpose $G_{\mathcal{P}}$ on top of $G_{\mathcal{D}}$ and notice that to each edge of $G_{\mathcal{D}}$ corresponds an edge of $G_{\mathcal{P}}$ which is rotated 90° clockwise. See Fig. 6. We couple the two lattices so that an edge in $G_{\mathcal{D}}$ is closed if and only if the corresponding edge in $G_{\mathcal{P}}$ is open. It is not hard to see that
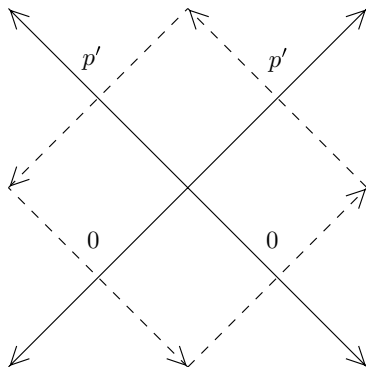
there is an open path from level 0 to level T in $G_{\mathcal{P}}$ if and only if there is no open path from the right boundary to the left boundary in $G_{\mathcal{D}}$. So it remains to compute an upper bound on the latter probability. Fix any two boundary nodes in $G_{\mathcal{D}}$, say $v_l = v_{1,\eta}$ and $v_r = v_{n',\zeta}$ for some η, ζ. The number of paths of length L between v_r and v_l is at most 3^L. Each such path makes $n' - 1$ more moves to the left than to the right. In particular, the number of moves to the left is at least $L/2$. Moreover, each edge going to the left has a probability $1 - p'$ of being open. So the probability that there is a path between v_r and v_l (which we denote $v_r \to v_l$) is at most

$$\mathbb{P}[v_r \to v_l] \leq \sum_{L=n'-1}^{+\infty} 3^L (1-p')^{L/2} \leq \frac{(3\,2^{-d/200})^{\frac{n-1}{2}}}{1 - 3\,2^{-d/200}},$$

for d large enough. There are at most T^2 pairs of boundary nodes so by the union bound

$$\mathbb{P}[s'_{i,T} = 0, \ \forall i \in \{2, 4, \ldots, n' - 1\}] \leq T^2 \frac{(3\,2^{-d/200})^{\frac{n-1}{2}}}{1 - 3\,2^{-d/200}} \leq 2^{-(d/1000)n},$$

for d large enough. $\square$

3 Win-stay lose-shift on graphs with large expansion

For this section, we consider the discrete-time version of the chain. That is, at every time step, we pick one edge uniformly at random and update the actions at the endpoints of that edge. Equivalently, we look at the discrete-time chain embedded in $\{\mathbf{A}_t\}_{t\geq 0}$ by stopping the chain every time a clock rings. Also, since we are looking for a lower bound on $T_{\mathbf{C}}$, we can speed up the chain by picking only those edges with at least one D endpoint. Denote the discrete-time sped-up chain $\{\mathbf{B}_k\}_{k\in\mathbb{N}}$.

The proof of Theorem 4 is based on the following geometric observation. Let U_k be the set of nodes defecting at time k and denote $N_k = |U_k|$. At the next update, N_k goes down by 2 if we pick an edge "inside" U_k and it goes up by 1 if we pick an edge on the "boundary" of U_k. Therefore, if the boundary of U_k is more than twice as big as the inside of U_k, on average the chain moves away from the fixed point $\mathbf{C}$.

Proof of Theorem 4: Let $U \subseteq V$ with $\alpha(n) \leq |U| \leq \beta(n)$. Note first that $\rho_{\alpha,\beta}(G) > 1/2 + \varepsilon$ implies

$$|E(U, U^c)| \geq \left(\frac{1}{2} + \varepsilon\right) \mathrm{vol}(U).$$

Let $\varepsilon' > 0$ such that $2 - \varepsilon' = (1/2 + \varepsilon)^{-1}$. Then

$$|E(U, U^c)| + 2|E(U, U)| = \mathrm{vol}(U) \leq (2 - \varepsilon')|E(U, U^c)|,$$

which implies

$$2|E(U, U)| \leq (1 - \varepsilon')|E(U, U^c)|.$$

Therefore there is an $\varepsilon'' > 0$ such that if $\alpha(n) \leq N_k \leq \beta(n)$, then

$$N_{k+1} = \begin{cases} N_k + 1, & \text{with probability at least } \frac{2}{3} + \varepsilon'', \\ N_k - 2, & \text{with probability at most } \frac{1}{3} - \varepsilon''. \end{cases}$$

Let

$$a = \left[\frac{1}{2}\left(\frac{2}{3} + \varepsilon''\right)\left(\frac{1}{3} - \varepsilon''\right)^{-1}\right]^{1/3} > 1.$$

It is easy to check that

$$\left(\frac{2}{3} + \varepsilon''\right)a^{-1} + \left(\frac{1}{3} - \varepsilon''\right)a^2 < 1.$$

Therefore,

$$W(N_k) = a^{n - N_k},$$

is a bounded nonnegative supermartingale on $\{\alpha(n) \leq N_k \leq \beta(n)\}$. Using the optional sampling theorem as in Dyer et al. (2002), it follows that the probability of N_k crossing the interval $[\alpha(n), \beta(n)]$ is less than $a^{-(\beta(n) - \alpha(n))}$ for n large enough. The theorem immediately follows.
□

4 Concluding remarks

The setup analyzed here may seem rather stylized. Although, as we mentioned in the Introduction, the study of nonreversible Markov chains such as IPD is very challenging and, in general, requires a case-by-case analysis. In particular, our results reinforce the general view that the convergence of nonreversible local dynamics is not determined by simple geometric properties of the underlying graph, such as expansion or treewidth. Therefore, it would be ill-advised to try and generalize our results to other similar local learning dynamics on the

basis of the analysis presented here. However, one could venture the following conjectures concerning the WSLS dynamics on bounded-degree graphs:

- **Conjecture 1**: There is a constant d such that for all n large enough and for all trees of minimum degree d with n nodes, the emergence of cooperation is exponentially slow in n.
- **Conjecture 2**: A stronger version of the previous conjecture is the following. There is a constant d such that for all n large enough and for all bounded-degree graphs of minimum degree d with n nodes, the emergence of cooperation is exponentially slow in n.
- **Conjecture 3**: For all d, there is a constant $l = l(d)$ such that for all n large enough and for all bounded-degree graphs of maximum degree d with n nodes, replacing all edges with paths of length at least l makes the emergence of cooperation polynomially fast (in the size of the resulting graph).

References

Aldous, D., & Fill, J. (2006). Reversible markov chains and random walks on graphs, Book in Preparation.

Axelrod, R. (1984). *The evolution of cooperation*. Basic Books.

Berger, N., Kenyon, C., Mossel, E., & Peres, Y. (2005). Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields*, *131*(3), 311–340. Extended abstract by Kenyon, Mossel and Peres appeared in Proceedings of 42nd IEEE Symposium on Foundations of Computer Science (FOCS) 2001, 568–578.

Durrett, R. (1984). Oriented percolation in two dimensions. *The Annals of Probability*, *12*, 999–1040.

Durrett, R. (1996). *Probability: theory and examples*. Duxbury.

Dyer, M., Goldberg, L. A., Greenhill, C., Istrate, G., & Jerrum, M. (2002). Convergence of the iterated prisoner's dilemma game. *Combinatorics, Probability, and Computing*, *11*, 135–147.

Friedman, J., Kahn, J., & Szemeredi, E. (1989). On the second eigenvalue in Random Regular Graphs. In *Proceedings of ACM STOC*.

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games*. MIT Press.

Kahale, N. (1995). Eigenvalues and expansion of regular graphs. *Journal of the ACM*, *42*, 1091–1106.

Kittock, J. E. (1995). Emergent conventions and the structure of multi-agent systems. In L., Nadel and D., Stein (Eds.), Vol. VI of *Santa Fe Institute Studies in the Sciences of Complexity Lecture*, Santa Fe Institute, *1993 lectures in complex systems: proceedings of the 1993 complex systems summer school*, Addison-Wesley.

Liggett, T. M. (1985). *Interacting particle systems*. Springer.

Liggett, T. M. (1999). *Stochastic interacting systems: contact, voter and exclusion processes*. Springer.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*, 56–58.

Shoham, Y., & Tennenholtz, M. (1993). Co-learning and the evolution of social activity. *Mimeo*.

Shoham, Y., & Tennenholtz, M. (1997). On the emergence of social conventions: Modelling, analysis and simulations. *Artificial Intelligence*, *94*, 139–166.