Check for updates

Adv. Appl. Probab. 1–47 (2026)
*doi:*10.1017/apr.2026.10074

LEARNING MODELS ON ROOTED REGULAR TREES WITH MAJORITY UPDATE POLICY: CONVERGENCE AND PHASE TRANSITION

MOUMANTI PODDER,* *Indian Institute of Science Education and Research Pune*
ANISH SARKAR,** *Indian Statistical Institute, Delhi Centre*

Abstract

We study a model of social learning on rooted regular trees. An agent is stationed at each vertex of $\mathbb{T}_m$, the rooted tree in which each vertex has precisely m children, and at any time step $t \in \mathbb{N}_0$, the agent is allowed to select one of two available technologies: B and R. Let the technology chosen by the agent at vertex v of $\mathbb{T}_m$, at time step t, be $C_t(v)$. We begin with the independent and identically distributed (i.i.d.) collection $\{C_0(v) : v \in \mathbb{T}_m\}$, where $C_0(v) = B$ with probability π_0. During the epoch t, the agent at vertex v performs an experiment that results in success with probability p_B if $C_t(v) = B$, and with probability p_R if $C_t(v) = R$. If the children of v are denoted $v_1, \ldots, v_m$, the agent at v updates their technology to $C_{t+1}(v) = B$ if the number of successes among all v_i (where $i \in \{1, 2, \ldots, m\}$) with $C_t(v_i) = B$ exceeds, strictly, the number of successes among all v_j (where $j \in \{1, 2, \ldots, m\}$) with $C_t(v_j) = R$. If these two numbers are equal then the agent at v sets $C_{t+1}(v) = B$ with probability $1/2$. In all other cases, $C_{t+1}(v) = R$. We show that $\{C_t(v) : v \in \mathbb{T}_m\}$ is i.i.d. as well, with $C_t(v) = B$ with probability π_t, where the sequence $\{\pi_t\}_{t\in\mathbb{N}_0}$ converges to a fixed point π, in $[0, 1]$, of a function g_m. We show that for $m \geqslant 3$, there exists a $p(m) \in (0, 1)$ such that g_m has the unique fixed point $1/2$ when $p \leqslant p(m)$, and three distinct fixed points, of the form α, $1/2$, and $1 - \alpha$, for some $\alpha \in [0, 1/2)$ when $p > p(m)$. When $m = 3$, $p_B = 1$, and $p_R \in [0, 1)$, we show that the function g_3 (i) has a unique fixed point, 1, when $p_R < \sqrt{3} - 1$, (ii) has two distinct fixed points, one of which is 1, when $p_R = \sqrt{3} - 1$, and (iii) has three distinct fixed points, one of which is 1, when $p_R > \sqrt{3} - 1$. When g_m has multiple fixed points, we also specify which of these fixed points π equals, depending on π_0. Finally, for $m = 2$, we describe the behaviour of g_2 for *all* values of p_B and p_R.

Keywords: Learning models; social learning; phase transitions; convergence of stochastic processes; interacting particle systems; rooted regular trees; diffusion of technologies

2020 Mathematics Subject Classification: Primary 68Q32
Secondary 91D30; 68T05; 91D10; 60K35; 68Q87

1. Introduction

1.1. Overview of our learning model

For any positive integer m with $m \geqslant 2$, let $\mathbb{T}_m$ indicate the infinite rooted tree, with root ϕ, in which each vertex has precisely m children. Let $V(\mathbb{T}_m)$ denote the set of all vertices of $\mathbb{T}_m$,

Received 14 May 2025 UTC; accepted 3 July 2026 UTC.
* Postal address: Department of Mathematics, Indian Institute of Science Education and Research Pune, Pune, India. Email address: moumanti@iiserpune.ac.in
** Postal address: Stat-Math Unit, Indian Statistical Institute, Delhi Centre, New Delhi, India. Email address: anish@isid.ac.in
© The Author(s), 2026. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

and imagine an agent occupying each vertex of $V(\mathbb{T}_m)$. We consider a *learning model* in which each agent, at each time step (where time is indexed by the set $\mathbb{N}_0$ of non-negative integers), has to adopt one of two possible technologies that we henceforth refer to as *colours* or *states* and denote by the letters B and R. The state of the agent at a vertex $v \in V(\mathbb{T}_m)$ at time step t, for each $t \in \mathbb{N}_0$, is denoted by $C_t(v)$. Conditioned on the states of the agents at all the vertices of $V(\mathbb{T}_m)$ at time step t, the (random) state $C_{t+1}(v)$ to which the agent at v updates itself at time step $t+1$ has a probability distribution that is a function of $C_t(v_1), C_t(v_2), \ldots, C_t(v_m)$, where $v_1, v_2, \ldots, v_m$ are the children of v.

Such a model can be viewed from a number of different perspectives, such as

(i) as a discrete-time *interacting system of particles* on $\mathbb{T}_m$ with infinitely many changes allowed to happen at each time step,

(ii) as a model for *social learning*, for understanding the *diffusion of technologies* throughout an infinite population of agents, and

(iii) as a *probabilistic finite state tree automaton* on $\mathbb{T}_m$.

We delve deeper into each of the above motivating reasons for the study of such models in Subsection 1.4.

Each model studied in this paper admits a crucial component known as a *policy function*: it governs how the agent at v, for each $v \in V(\mathbb{T}_m)$, *learns* from the states $C_t(v_i)$, $i \in \{1, 2, \ldots, m\}$, of the agents situated at the children $v_1, v_2, \ldots, v_m$ of v at time step t, and subsequently updates its own state to $C_{t+1}(v)$ at time step $t+1$. When such a model is viewed as a *probabilistic tree automaton* (PTA), as explained in Subsection 1.4, the policy function is referred to as the *stochastic update rule* associated with that PTA. We are concerned with studying the *absolute majority policy* function, which we describe in detail in Subsection 1.2.

1.2. Absolute majority policy

In this model, the agent occupying vertex v, for each $v \in V(\mathbb{T}_m)$, is endowed with two random variables, $X_t(v)$ and $Y_t(v)$, for each time step $t \in \mathbb{N}_0$. For each $v \in V(\mathbb{T}_m)$ and each $t \in \mathbb{N}_0$, the random variable $X_t(v)$ captures the *outcome* of the (random) experiment performed by the agent located at v, during the time step or *epoch* t, using the technology $C_t(v)$ that this agent is currently equipped with. This experiment has two possible outcomes—*success* and *failure*—so $X_t(v)$ is a Bernoulli random variable that acts as an indicator for the event of success. We let $X_t(v)$ follow Bernoulli(p_B) conditioned on $C_t(v) = B$, and we let $X_t(v)$ follow Bernoulli(p_R) conditioned on $C_t(v) = R$, where $p_B, p_R \in [0, 1]$ are preassigned parameters in our model. The random variable $Y_t(v)$ follows Bernoulli$\left(\frac{1}{2}\right)$ for each $v \in V(\mathbb{T}_m)$ and each $t \in \mathbb{N}_0$. It is assumed that the entire collection $\{X_t(v) : v \in V(\mathbb{T}_m)\} \cup \{Y_t(v) : v \in V(\mathbb{T}_m)\}$ of random variables is independent for *each* time step $t \in \mathbb{N}_0$, and the collection $\{Y_t(v) : v \in V(\mathbb{T}_m)\}$ is independent of $\{C_s(u) : u \in V(\mathbb{T}_m), s \in \{0, 1, \ldots, t\}\}$ for each $t \in \mathbb{N}_0$.

Conditioned on $\{C_t(u) : u \in V(\mathbb{T}_m)\}$, the agent at v updates its state to $C_{t+1}(v)$ according to the following rule, where $v_1, v_2, \ldots, v_m$ denote the children of v:

$$C_{t+1}(v) = \begin{cases} B & \text{if } \sum_{i=1}^m X_t(v_i)\mathbf{1}_{C_t(v_i)=B} > \sum_{i=1}^m X_t(v_i)\mathbf{1}_{C_t(v_i)=R}, \\ B & \text{if } \sum_{i=1}^m X_t(v_i)\mathbf{1}_{C_t(v_i)=B} = \sum_{i=1}^m X_t(v_i)\mathbf{1}_{C_t(v_i)=R} \text{ and } Y_t(v) = 1, \\ R & \text{otherwise.} \end{cases} \tag{1}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

In words, this can be described as follows: the agent at the vertex v counts the number $\sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=B}$ of successful experiments performed, during epoch t, by all those agents that occupy the children of v *and* that are in state B at time step t, and likewise, it also counts the number $\sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=R}$ of successful experiments performed, during epoch t, by all those agents that occupy the children of v *and* that are in state R at time step t. If the former count strictly exceeds the latter then the agent at v updates its state to B at time step $t+1$, while if the latter strictly exceeds the former then the agent at v updates its state to R at time step $t+1$. If the two counts are equal then the agent at v tosses a fair coin in order to break the tie. If the coin lands as a head, indicated by $Y_t(v)=1$, then the agent at v updates its state to B at time step $t+1$, and if the coin lands as a tail, it updates its state to R at time step $t+1$.

We begin the process at time step $t=0$ with an independent and identically distributed (i.i.d.) assignment of states to the vertices of $\mathbb{T}_m$, i.e. the collection $\{C_0(v) : v \in V(\mathbb{T}_m)\}$ is i.i.d. with

$$C_0(v) = B \text{ with probability } \pi_0 \text{ and } C_0(v) = R \text{ with probability } 1-\pi_0 \tag{2}$$

for each $v \in V(\mathbb{T}_m)$. Letting ν_t denote the joint law of $\{C_t(v) : v \in V(\mathbb{T}_m)\}$ for each $t \in \mathbb{N}_0$, we are concerned in this paper with answering the following question: does $\{\nu_t\}_{t\in\mathbb{N}_0}$ converge as $t \to \infty$, and if it does, what is its distributional limit? We investigate this question for *all* values of $(p_B, p_R) \in [0,1]^2$ when $m=2$, whereas when $m \geqslant 3$, we consider two *regimes* of values of the parameter pair (p_B, p_R). The first of these regimes is where $p_B = p_R = p$ for some $p \in [0,1]$, i.e. an agent, located at any vertex of $\mathbb{T}_m$, stands the same chance of achieving a successful outcome in the experiment performed during epoch t irrespective of whether its state at time step t is B or R. We present a full description of what happens in such a scenario as the common value p of the two parameters is allowed to vary over $[0,1]$. The complement of this regime is where $p_B \neq p_R$, but when $m \geqslant 3$, the analysis for $p_B \neq p_R$, without additional assumptions, quickly becomes intractable, and it becomes difficult to conclude anything about the distributional limit(s) of $\{\nu_t\}_{t\in\mathbb{N}_0}$ (depending on the initial distribution ν_0 given by (2), and on the values of p_B and p_R). We, therefore, focus on the special case of $p_B = 1$ and $p_R \in [0,1)$, and we demonstrate, via the relatively less unwieldy examples of $m=2$ and $m=3$, that the conclusion is rather non-intuitive and further investigations are warranted. This special case can be considered 'extremal' in the sense that a vertex that is in state B is *sure* to achieve success when it performs its experiment (using technology B), since $p_B = 1$, whereas the experiment performed (with technology R) by a vertex in state R has a positive probability of resulting in a failure. Intuition seems to suggest that in such a scenario, eventually, every vertex of $\mathbb{T}_m$ would opt for the (evidently superior) technology B, but it turns out (see Section 5) that this is not necessarily the case, and the distributional limit depends heavily on the value of p_R.

1.3. Main results

We begin with a result that is applicable to *all* values of the parameter pair (p_B, p_R).

Theorem 1. *Recall that ν_t indicates the joint law of $\{C_t(v) : v \in V(\mathbb{T}_m)\}$ for each $t \in \mathbb{N}_0$, and recall that we begin from an i.i.d. assignment of states (from $\{B, R\}$) to the vertices of $\mathbb{T}_m$, as described in (2). The following are true.*

(i) *For each $t \in \mathbb{N}$, the random variables $C_t(v)$, for $v \in V(\mathbb{T}_m)$, are i.i.d. with*

$$C_t(v) = B \text{ with probability } \pi_t \text{ and } C_t(v) = R \text{ with probability } 1-\pi_t \tag{3}$$

for each $v \in V(\mathbb{T}_m)$.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

(ii) *The sequence* $\{\nu_t\}_{t\in\mathbb{N}_0}$ *converges if and only if the sequence* $\{\pi_t\}_{t\in\mathbb{N}_0}$ *converges, as* $t\to\infty$, *and in that case, if* $\pi=\lim_{t\to\infty}\pi_t$ *then* $\{\nu_t\}_{t\in\mathbb{N}_0}$ *converges to* ν, *where* ν *is the joint law of the collection* $\{C_\infty(v): v\in V(\mathbb{T}_m)\}$ *of i.i.d. random variables* $C_\infty(v)$, *where*

$$C_\infty(v)=B \text{ with probability } \pi \text{ and } C_\infty(v)=R \text{ with probability } 1-\pi \tag{4}$$

for each $v\in V(\mathbb{T}_m)$.

(iii) *The limit* $\pi=\lim_{t\to\infty}\pi_t$, *when it exists, is a fixed point of the function* $g_m:[0,1]\to[0,1]$ *defined as*

$$g_m(x)=\sum_{k=0}^{m} f_m(k)\binom{m}{k}x^k(1-x)^{m-k}, \tag{5}$$

in which

$$f_m(k)=\mathbb{P}\left[A_k>B_{m-k}\right]+\frac{1}{2}\,\mathbb{P}\left[A_k=B_{m-k}\right], \tag{6}$$

where the random variable A_k *follows Binomial*(k,p_B), *the random variable* B_{m-k} *follows Binomial*$(m-k,p_R)$, *and* A_k *and* B_{m-k} *are independent of each other.*

From the last assertion made in the statement of Theorem 1, it becomes apparent that we need to answer the following questions.

(i) Does the function g_m have a unique fixed point in $[0,1]$?

(ii) If g_m has multiple fixed points in $[0,1]$, which of these are attractive and which are repulsive?

Let us now make the following simple observation.

Remark 1. When g_m has a unique fixed point, say π, in $[0,1]$, it must be attractive, i.e. no matter what value $\pi_0\in[0,1]$ we choose in (2), the sequence of laws $\{\nu_t\}_{t\in\mathbb{N}_0}$ converges, as $t\to\infty$, to ν defined by (4).

Proof. Let us consider any subsequence $\left\{\pi_{t_k}\right\}_{k\in\mathbb{N}}$ of the sequence $\{\pi_t\}_{t\in\mathbb{N}_0}$ (defined in (3)). Since this subsequence is bounded, it has a further subsequence, say $\left\{\pi_{t_{k_\ell}}\right\}_{\ell\in\mathbb{N}}$, that converges as $\ell\to\infty$. From 1 of Theorem 1, we know that $\lim_{\ell\to\infty}\pi_{t_{k_\ell}}$ has to be a fixed point of g_m, and as g_m has a unique fixed point, π, in $[0,1]$, we must have $\lim_{\ell\to\infty}\pi_{t_{k_\ell}}=\pi$. This allows us to conclude that the entire sequence $\{\pi_t\}_{t\in\mathbb{N}_0}$ converges to π as $t\to\infty$, and the rest of the conclusion in Remark 1 follows from 1 of Theorem 1. □

Before moving on to higher values of m, we state the following result concerning $m=2$ and *all* possible values of p_B and p_R.

Theorem 2. *The function* g_2 *is strictly convex for* $p_R>p_B$, *strictly concave for* $p_R<p_B$, *and linear for* $p_B=p_R$. *In each of these cases, except when* $p_B=p_R=1$, *the function* g_2 *has a unique fixed point in* $(0,1)$. *When* $p_B=p_R=1$, *the entire interval* $[0,1]$ *constitutes the set of fixed points of* g_2.

From Remark 1, it is evident that for $m=2$ and when $(p_B,p_R)\neq(1,1)$, the sequence $\{\nu_t\}_{t\in\mathbb{N}_0}$ converges to ν, as $t\to\infty$, where ν is as defined by (4), with π being the unique fixed point of g_2. For $m=2$ and $p_B=p_R=1$, for each $\pi_0\in[0,1]$ (where π_0 is as defined in (2)), the law ν_t is the same as ν_0 for every $t\in\mathbb{N}$, so the limit equals ν_0 itself.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

We now come to the two specific regimes of values of the parameter pair (p_B, p_R) that, as mentioned towards the end of Subsection 1.2, are considered in this paper. When $p_B = p_R = p$ for some $p \in [0, 1]$, the learning model with the absolute majority policy described in Subsection 1.2 admits the following phase transition phenomenon.

Theorem 3. *Recall the function g_m defined in (5), and consider $p_B = p_R = p$ in the definition of the function f_m in (6). For each $m \in \mathbb{N}$ with $m \geqslant 2$, there exists $p(m) \in (0, 1)$ such that*

(i) *for all $p \leqslant p(m)$, the function g_m has a unique fixed point in $[0, 1]$, which is $1/2$;*

(ii) *for all $p > p(m)$, the function g_m has precisely three fixed points in $[0, 1]$, of the form α, $1/2$, and $1 - \alpha$, where $\alpha = \alpha(p, m) < 1/2$ is some function of p and m.*

Recall π_t from (3), for each $t \in \mathbb{N}$. For every choice of the initial distribution (i.e. for every choice of π_0, as defined in (2)), the sequence $\{\pi_t\}_{t \in \mathbb{N}_0}$ converges, as $t \to \infty$, to π (as defined in the statement of Theorem 1). Here, $\pi = 1/2$ for each $\pi_0 \in [0, 1]$ when $p \leqslant p(m)$, whereas when $p > p(m)$, we have (i) $\pi = \alpha$ when $\pi_0 \in [0, 1/2)$, (ii) $\pi = 1/2$ when $\pi_0 = 1/2$, and (iii) $\pi = 1 - \alpha$ when $\pi_0 \in (1/2, 1]$.

Remark 2. Note that in the regime where $p \leqslant p(m)$, since g_m has a unique fixed point in $[0, 1]$, and this equals $1/2$, by Remark 1 we already know that $\pi = 1/2$ for each $\pi_0 \in [0, 1]$. Therefore, in the last assertion made in the statement of Theorem 3, we need focus only on the regime where $p > p_m$.

In particular, we show, via analytical methods in Subsection 4.1, that for $m = 3$, the threshold $p(m) \approx 0.557507$, whereas for $m = 4$, the threshold $p(m) \approx 0.42842$.

When $p_B = 1$ and $p_R \in [0, 1)$, our learning model is devoid of any phase transition for $m = 2$ (which follows as a special case of Theorem 2), whereas a striking and rather non-intuitive phenomenon of phase transition is observed when $m = 3$, as stated below.

Theorem 4. *When $m = 3$, $p_B = 1$, and $p_R \in [0, 1]$, the function g_m defined in (5):*

(i) *has a unique fixed point in $[0, 1]$, which is 1, when $p_R \in [0, \sqrt{3} - 1)$;*

(ii) *has two distinct fixed points in $[0, 1]$, namely $\alpha = 2/3 - 1/\sqrt{3}$ and 1, when $p_R = \sqrt{3} - 1$; and*

(iii) *has three distinct fixed points in $[0, 1]$, one of which is 1, when $p_R \in (\sqrt{3} - 1, 1]$.*

Let π_0 be as defined in (2), and recall $\pi = \lim_{n \to \infty} \pi_n$, where π_n is as defined in (3). When $p_R \in [0, \sqrt{3} - 1)$, we have $\pi = 1$. When $p_R = \sqrt{3} - 1$, we have $\pi = \alpha$ when $\pi_0 \in [0, \alpha]$ and $\pi = 1$ when $\pi_0 \in (\alpha, 1]$. When $p_R \in (\sqrt{3} - 1, 1)$, let the two fixed points other than 1 be denoted by $\alpha_1 = \alpha_1(p_R)$ and $\alpha_2 = \alpha_2(p_R)$, with $0 < \alpha_1 < \alpha_2 < 1$. In this case, we have $\pi = \alpha_1$ whenever $\pi_0 \in [0, \alpha_2)$, and $\pi = 1$ whenever $\pi_0 \in (\alpha_2, 1]$, whereas $\pi = \alpha_2$ when $\pi_0 = \alpha_2$.

The reason we describe this result as non-intuitive is as follows. The value $p_B = 1$ ensures that *every* experiment performed with the technology labelled B is bound to result in a success. Evidently, the technology labelled B is superior to the technology labelled R, unless $p_R = 1$. This would seem to suggest that *eventually* the state of each vertex of $\mathbb{T}_m$ would be updated to B and would remain that way, but this is not the case when p_R is sufficiently large (i.e. for $p_R \geqslant \sqrt{3} - 1$). The phase transition exhibited by our model in Theorem 4, therefore, seems unusual.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

1.4. Motivation and a brief discussion of relevant literature

As alluded to in Subsection 1.1, there is a myriad of perspectives from which the study of our model can be motivated. For any two vertices $u, v \in V(\mathbb{T}_m)$, with u the parent of v, let us refer to the agent located at v as an *influencing neighbour* to the agent located at u. Our model can then be viewed as an interacting particle system in which each agent, at each time step, is allowed to update its policy or protocol on some matter by taking note of the policies or protocols adopted, and the outcomes (such as rewards or penalties) obtained as a consequence, by its influencing neighbours at the *previous* time step. Each agent can adopt one of two available policies (such as a *liberal* stance and a *conservative* stance) at the beginning of time step t, following which it executes a plan of action (as decreed by the adopted policy) *during* the epoch t. Such a plan of action (which is a random experiment) has two possible outcomes: a *desirable* one and an *undesirable* one. At the end of time step t, each agent counts the number of its influencing neighbours who adopted the liberal stance at the beginning of time step t and who were then rewarded with the desirable outcome, and the number of its influencing neighbours who adopted the conservative stance at the beginning of time step t and who were then rewarded with the desirable outcome. If the former exceeds the latter, the agent adopts the liberal stance at the beginning of time step $t+1$, whereas if the latter exceeds the former, the agent adopts the conservative stance at the beginning of time step $t+1$. If the two counts are equal, the agent decides which stance to adopt by tossing a fair coin and thereby breaking the tie.

The interpretation presented above is reminiscent of the well-known and well-studied *voter model*. In the classical *linear voter model* (see, for instance, [57], which introduced the voter model as a *continuous-time proximity process*, as well as [68], [67], and [66] for a detailed review of old and new developments in this topic of research), individuals or agents occupy the elements (referred to as *sites*) of an arbitrary countable set S, and each agent, at any given point in time, can have one of two possible opinions on an issue. For each site $x \in S$, at exponential times with a specified rate, the agent located at x chooses a site $y \in S \setminus \{x\}$ with probability $p(x, y)$ (where $p(x, y) \geqslant 0$ for each $y \in S \setminus \{x\}$ and $\sum_{y \in S \setminus \{x\}} p(x, y) = 1$) and adopts the opinion being currently held by the agent at y.

Yet another interpretation of the voter model is that of two species, fairly matched, competing for territory, and the invasion by one species of the territory belonging to the other. In our model, each vertex v of $\mathbb{T}_m$ can be viewed as a city, and B and R can be thought of as two belligerent armies contesting the city. At any time step t, the event $C_t(v) = B$ indicates that army B retains dominion over the city v at the *beginning* of time step t, but that there are ongoing clashes between B and R *during* the epoch t. Conditioned on $C_t(v) = B$, the event $X_t(v) = 1$ indicates that the *defenders*, i.e. the armed forces of B, are able to defeat the *invaders*, i.e. the armed forces of R, in the battle that takes place in v during the epoch t. It is not only reasonable, but is even practical, to assume that, in the event of $X_t(v) = 0$, even though the invaders clinch a victory over the defenders, they incur heavy losses, and those among them that remain alive are not able to travel from v to anywhere else for further attempts at conquests. A similar interpretation is true if $C_t(v) = R$. Letting $v_1, v_2, \ldots, v_m$ denote the children of v, we may imagine them to be, in some sense, 'border cities' or 'outposts' of the 'main city' v, with routes leading *to* v (along which armies can travel). From each outpost v_i where the defenders B defeat the invaders R in the battle that takes place during the epoch t, a battalion of soldiers loyal to B is dispatched to v by the end of the epoch t. Likewise, from each v_i where the defenders R defeat the invaders B in the battle that takes place during the epoch

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

t, a battalion of soldiers loyal to R is dispatched to v by the end of the epoch t. The inequality $\sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=B} > \sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=R}$ then indicates that the soldiers who are loyal to B and who travel to v from its outposts outnumber the soldiers who are loyal to R and who travel to v from its outposts, resulting in v becoming occupied by B by the beginning of the epoch $t+1$. On the other hand, when $\sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=B} = \sum_{i=1}^{m} X_t(v_i)\mathbf{1}_{C_t(v_i)=R}$, the number of soldiers loyal to B who travel to v from its outposts equals the number of soldiers loyal to R who travel to v from its outposts, and in this case, v is equally likely to be conquered by either of the armies by the beginning of the epoch $t+1$. In this regard, we mention [26], in which a simple stochastic model is devised for two fairly matched species competing for territory along the boundary or frontier of their respective regions of dominion, wherein the result of the conflict is the invasion by one of the species of territory previously held by the other. Asymptotic results for the pattern of territories held by a single species are presented in [26], along with the exhibition of markedly distinct behaviour between one-dimensional and two-dimensional competitions on one hand, and three-dimensional competition on the other.

Interacting particle systems are indispensable for modelling complex phenomena, in natural and social sciences, that involve a very large number of interrelated components (such as the flow of traffic on highways, interaction among the constituents of a cell, opinion dynamics, the spread of epidemics or fires, reaction diffusion systems, crystal surface growth, and chemotaxis). For instance, the *biased voter model*, a generalization of the voter model described above, has been studied in [101], [19], [20], and [61], in the context of how cancerous cells spread and metastasize throughout the body of an organism. In [80], *binary opinion dynamics* in a fully connected network of interacting agents is studied, with agents being biased towards one of the two opinions and allowed to interact according to (i) the voter rule, where an updating agent simply copies the opinion of a different, randomly sampled, agent, and (ii) the majority rule, where an updating agent samples multiple other agents and adopts the opinion that is in majority among the members of this selected sample. The diverse applications of the voter model and its variants in the field of social sciences, especially in the analysis of discrete and continuous opinion dynamics, can be seen, for example, in [49], [75], [76], [77], [29], [91], [90], [72], and [23].

We mention here that the two other extensively studied models of interacting particle systems are the *contact processes* (see, for instance, [50], [36], [16], [92, 93] and [67]) and the *exclusion processes* (see, for instance, [65], [30], [85], [87], [88], [86], [94], [71], and [67]). Of these, the contact process finds applications in the study of the *susceptible-infected-susceptible* (SIS) model of *compartmental epidemiology* (see, for instance, [14]), wherein every vertex of the network under consideration, at any point in time, is either *infected* or *healthy but susceptible*, and whereas each infected vertex becomes healthy, independent of all else, at a given constant rate, a healthy vertex becomes infected at a rate that equals the product of the propagation ratio of the disease and the number of infected neighbours of that vertex (at that time). Much like the model considered in this paper, this, too, is a model involving agents stationed at the vertices of a given (or random) graph or network (in particular, Berger *et al.* [14] studied the contact process, emulating the spread of a virus throughout a network of computers in the presence of a particular class of antivirus software, on scale-free graphs in preferential attachment models), wherein each agent can be in one of two possible states at any given point in time, and if an agent is currently healthy, the rate at which it transitions into the infected state is influenced by the number of infected neighbours it currently possesses. In [37], the contact process/SIS model is studied on *small-world networks* proposed by Newman and Watts [83], wherein all pairs of vertices that are a distance m or less apart are connected (referred to as

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

short-range connections), and in addition, there is a density p of edges (called *long-range connections*) that connect vertices to neighbours that are farther away and are chosen at random from the graph.

The model studied in this paper also seems to be redolent of models for *competing first passage percolation*, such as that studied in [4]. Here, at any time $t \geqslant 0$ of the stochastic process, the set of vertices V of a given graph $G = (V, E)$ can be partitioned into three subsets, namely (i) $\mathcal{B}_t$, consisting of all the vertices painted blue so far, (ii) $\mathcal{R}_t$, consisting of all the vertices painted red so far, and (iii) $V \setminus (\mathcal{B}_t \cup \mathcal{R}_t)$, consisting of the yet-uncoloured vertices, and the Markov process evolves by having each yet-uncoloured vertex v of V assigned (i) the colour blue with rate equal to $\beta\,|N(v) \cap \mathcal{B}_t|$ and (ii) the colour red with rate equal to $\rho\,|N(v) \cap \mathcal{R}_t|$, where $N(v)$ indicates the set of all neighbours of v in G, and ρ and β are positive *rates* that remain fixed throughout the process. Here, however, once a vertex has been coloured, it retains that colour forever in the future. Continuing with this line of thought, we are ushered into the world of *two-type Richardson models*, introduced in [52], wherein two infections, type 1 and type 2, starting from two distinct sources, s_1 and s_2, respectively, in the cubic grid $\mathbb{Z}^d$, compete to invade the sites of $\mathbb{Z}^d$, each progressing like a first-passage percolation process on $\mathbb{Z}^d$ governed by the same family of i.i.d. exponential random variables, but with the assumption that once a site is captured by the type i infection, it remains of type i forever and cannot transmit the other infection. In [53], the two-type Richardson model on $\mathbb{Z}^d$, with $d \geqslant 2$, is studied in the asymmetric case with different infection rates: a site $\mathbf{x}$ in $\mathbb{Z}^d$ that is yet uninfected may either become type 1 infected at a rate that is equal to λ_1 times the number of type-1-infected neighbours of $\mathbf{x}$ or else become type 2 infected at a rate that is equal to λ_2 times the number of type-2-infected neighbours of $\mathbf{x}$, with $\lambda_1 \neq \lambda_2$. The main question of interest in both [52] and [53] is whether it is possible, with positive probability, for both the set of sites that are type 1 infected and the set of sites that are type 2 infected to keep growing indefinitely as time approaches ∞ (in other words, whether *coexistence* is possible). In [43], it is shown that coexistence occurs with strictly positive probability when stationary ergodic first-passage percolation is considered (on $\mathbb{Z}^d$ for $d \geqslant 2$, and also on an infinite cluster of Bernoulli percolation) and some additional assumptions (such as integrability conditions on the passage times, finite energy properties, and positivity conditions on the functional giving the directional asymptotic speeds) are imposed.

In [56], results from the study of stationary first-passage percolation are applied to the analysis of limiting configurations in a large family of *competing spatial growth models* that may be interpreted as variants of the two-type Richardson model, on $\mathbb{Z}^d$. In [31], a generalized version of the two-type Richardson model (with infection rates of the type 1 and type 2 infections being λ_1 and λ_2, respectively) is considered, in which it is assumed that at the start of the process, all sites in the set ξ_i are infected with type i, where ξ_1 and ξ_2 are two finite, disjoint subsets of $\mathbb{Z}^d$, and all sites in $\mathbb{Z}^d \setminus (\xi_1 \cup \xi_2)$ are uninfected. The set ξ_i is said to *strangle* the set ξ_j, for $\{i, j\} = \{1, 2\}$, if there exists no infinite self-avoiding path in $\mathbb{Z}^d$ that starts at a vertex in ξ_j and does not intersect ξ_i, and the pair (ξ_1, ξ_2) is said to be fertile if neither of them strangles the other. In [31], it is shown that as long as (ξ_1, ξ_2) is fertile, the actual choice of ξ_1 and ξ_2 has no impact on whether the probability of coexistence is strictly positive or not. In [32], the two-type Richardson model is studied on graphs other than the cubic lattice $\mathbb{Z}^d$, and the existence of graphs that (i) exhibit coexistence at some value of the parameter pair (λ_1, λ_2) such that $\lambda_1 \neq \lambda_2$ and (ii) exhibit non-coexistence when the ratio λ_1/λ_2 is brought closer to 1 is proved. In [33], unbounded sets of initially infected sites are considered, and it is shown that (i) if, initially, all sites of $\mathcal{H} \setminus (0, 0, \ldots, 0)$, where $\mathcal{H} = \{(x_1, x_2, \ldots, x_d) \in \mathbb{Z}^d : x_1 = 0\}$, are

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

type 1 infected and the site $(0, 0, \ldots, 0)$ is type 2 infected, then there is a positive probability for the type 2 infection to grow unboundedly if and only if its intensity strictly exceeds that of the type 1 infection (i.e. $\lambda_2 > \lambda_1$), whereas (ii) if, initially, all sites of $\mathcal{L} \setminus (0, 0, \ldots, 0)$, where $\mathcal{L} = \{(x_1, x_2, \ldots, x_d) : x_1 \leqslant 0, x_i = 0 \text{ for each } i \in \{2, \ldots, d\}\}$, are type 1 infected and the site $(0, 0, \ldots, 0)$ is type 2 infected, then there is a positive probability for the type 2 infection to grow unboundedly even if $\lambda_1 = \lambda_2$. In [44], a weakened version of non-coexistence for the two-type Richardson model with distinct infection rates for the two infection types is proved, i.e. assuming that the two types of infection propagate according to passage times that are distinct and stochastically comparable, [44] shows that when the weaker infection type survives, the space occupied by the stronger infection is almost impossible to detect, and in particular, in dimension 2, one of the two infection types occupies a set of sites with full density, while the other occupies a set of sites with null density.

In [34], the two-type Richardson model is studied on random graphs generated by the configuration model, and it is shown that, as the number of vertices tends to infinity, (i) if the degree distribution is a power law with exponent in (2, 3), one of the infection types will occupy all but finitely many vertices, and (2) which infection type wins is random, and either infection type has a positive probability of winning, irrespective of the values of λ_1 and λ_2. In the same thread, we mention [11], [98], and [97], which study the two-type Richardson model on scale-free random graphs with degree distributions satisfying asymptotic power laws with infinite variance, where the spread of each of the two types of infection is governed by i.i.d. passage times (whose distributions may depend on the type of infection considered). In [3], the two-type Richardson model is studied in two settings, namely a planar growth model related to two-neighbour bootstrap percolation, and a system of urns with graph-based interactions. In [2], the two-type Richardson model is studied on configuration models with given degree sequences, and as the total number of vertices approaches infinity, it is shown that (i) when $\lambda_1 = \lambda_2$ and the degree of a randomly chosen vertex has a finite second moment, the proportion of vertices infected by type 1 converges to a continuous random variable $V \in (0, 1)$, and (ii) when $\lambda_1 > \lambda_2$, the type 1 infection occupies all but a vanishing fraction of vertices. In [69], competition between two growth models with long-range interactions on a torus of size n in dimension d is studied.

We mention here [6], which admits, as a special case of the *influence model* studied therein, the *binary influence model*. Here, at any given point in time, each site of the network under consideration can be in one of two statuses, and transitions between these statuses at any given site can occur under the influence only of the neighbours (of that site) that have the opposite status (i.e. opposite the current status of that site). The binary influence model evolves in discrete time, and the structure of the underlying network plays a crucial role in determining whether all sites, eventually, reach a consensus status.

Some discussion is now in order as to the apparent connection that our work bears with [5], which studies a model that couples the growth of a random network with the nodal dynamics taking place on that network. To elucidate, [5] studies a dynamically evolving network, $\{G_n\}_n$, consisting of vertices coloured either blue or red. This begins with an initial graph G_0, with an assignment $\{C(u) : u \in G_0\}$ of colours to the vertices of G_0. At time stamp $(n + 1)$, a new vertex, say v_{n+1}, with m edges attached to it, is introduced. Each of these m edges, independent of all else, attaches to vertex u in G_n with probability proportional to $\deg_n(u)$, which is the degree of u in G_n. Once this step has been completed, the probability with which v_{n+1} is assigned the colour blue is equal to p_k, where k is the number of edges that connect v_{n+1} with blue vertices in G_n. The goal of [5] is to investigate the limiting behaviour of $a_n = (A_n + B_n)^{-1}A_n$,

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

where A_n is the number of blue vertices in G_n, and B_n is the number of red vertices in G_n. This is accomplished by studying the Markov process $\{(A_n, X_n)\}_n$, where X_n equals the sum of the degrees of all blue vertices in G_n, and the analysis is done by expressing $\{(A_n, X_n)\}_n$ as a *stochastic approximation* process. The conditional expectation of the increment $A_{n+1} - A_n$, conditioned on $\mathcal{F}_n$ (the σ-field containing all information regarding the stochastic process up to and including time stamp n) is computed, yielding a *drift function* (which takes as the argument the proportion $x_n = (X_n + Y_n)^{-1}X_n$, where X_n is as defined above, and Y_n is the sum of the degrees of all red vertices in G_n) that is closely connected with the function g_m defined in (5) (when $p_k = f_m(k) + k/m$ for each k). This similarity arises from the interpretation that each new vertex being introduced in the model studied in [5] has, as its 'children', the existing vertices to which it becomes connected via its m edges—of course, there may not, in this case, be m *distinct* 'children' of v_{n+1}, but the decision as to which colour v_{n+1} will adopt is made via m *distinct edges*. This is reminiscent of how, in our model, at the beginning of each epoch, the agent located at any vertex v of $\mathbb{T}_m$ relies on the feedback it receives from the agents located at the m children of v (over the previous epoch) in order to update its selection of technology. It should be noted, however, that the model presented in [5] differs from the model studied in this paper in two crucial aspects.

(i) The underlying network in [5] evolves with time, whereas the underlying network (i.e. $\mathbb{T}_m$) in our model remains fixed throughout.

(ii) In [5], the colour of each vertex, once chosen, remains fixed throughout, whereas in our model, *every* vertex gets the opportunity of updating its choice of which technology to adopt during *every* epoch.

Our analysis relies on the deduction and resolution of *recursive distributional equations* (see (10)), and since we are, ultimately, interested in the limiting joint distribution of the technologies adopted by all the agents on $\mathbb{T}_m$, i.e. the limit of the joint distribution of $(C_t(v) : v \in V(\mathbb{T}_m))$ as $t \to \infty$, recursive distributional equations seem to be the most direct and effective approach to our problem. On the other hand, recurrence relations of the forms given by equations (2) and (3) in [5] seem to be the most obvious approach to analysing the dynamics of the set-up studied in [5], following which it is natural to study these recurrence relations as a stochastic approximation process. While (1) does relate $C_{t+1}(v)$, for each v and each $t \in \mathbb{N}_0$, with $C_t(v_i)$ and $X_t(v_i)$ for $i \in \{1, 2, \ldots, m\}$, where $v_1, \ldots, v_m$ are the children of v, it does not take the same form as a stochastic approximation process, since we cannot get a recurrence relation directly connecting the random variables $C_{t+1}(v)$ and $C_t(v)$. Having said this, we speculate that when studying similar models of social learning involving (possibly) more complicated update policies, it may be the case that the sequence of random variables $\{C_t(v) : t \in \mathbb{N}_0\}$ for each $v \in V(\mathbb{T}_m)$ (or perhaps some other, more suitable, statistic, as a function of time) can be expressed as a stochastic approximation process, and from there onward, the techniques used in [5] may prove to be very helpful.

As a sequel to [5], we mention here [54], wherein a conjecture, postulating that an extension of the model studied in [5] to *more than two* opinions exhibits the same behaviour as its two-opinion counterpart (in that there is almost sure convergence to a limit which is a stationary point of a multidimensional vector field), is shown to be false. In the counterexample proposed in [54], each agent/vertex may adopt one of three available opinions, and the opinion-assignment mechanism (for the incoming vertex) is inspired by the game of 'rock-paper-scissors'. Another work closely related to [5] is [59], wherein the class of attachment

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

models studied in [5] is extended in order to take into account the possibility that the two types (or colours, i.e. red and blue, as mentioned in the previous paragraph) have different *fitnesses* (for instance, one of the types may be favoured, or seen as having higher fitness than the other), inspired by the *multiplicative fitness model* (see [18]) as well as the *additive fitness model* (see [1]). Finally, in [55], the work accomplished in [5] is extended to take into account a community structure in the network, and whether the proportions of vertices of each type synchronize between communities has been investigated.

Our learning model is also relevant in the context of *social learning*, a term that was first coined in [39]. In [39], each economic agent decides between two technologies, whose relative profitability is unknown, by taking into account the experiences of its neighbours and via exogenously specified rules of thumb that ignore historical data. Two learning environments are considered in [39], in one of which the same technology is considered optimal for *all* agents, whereas in the other, each technology is better for *some* of the agents. In [40], economic agents rely on information obtained via word-of-mouth communication in order to make decisions, without knowing the costs and benefits of the available choices. Two scenarios are considered in [40], the first of which allows a choice between two competing products with unequal qualities or pay-offs, while the second allows a choice between two products that are equally good. In [7], when the pay-offs from the various possible actions are unknown, agents use their own past experience as well as the experiences of agents belonging to their *neighbourhoods* to guide them in their decision-making, and a general framework to study the relationship between the structure of these neighbourhoods and the process of social learning is proposed and analysed. In [8], an approach to network formation is presented assuming that the link formed by one agent with another allows the former access, in part and in due course, to the benefits available to the latter via their own links, and that the cost of formation of a link is incurred only by the agent who initiates the formation, thereby allowing the network formation process to be formulated as a non-cooperative game. In [99], the process of network formation in a dynamic framework is analysed, with agents being allowed to form and sever links with one another, while in [73], agents are assumed to meet sequentially at random, myopically updating their links. More about the topic of social learning, as well as about learning models based on networks of agents, can be found, in, for example, [21], [45], [46], [10], [41], [47], [48], and [78].

As is evident from the description presented in Subsection 1.2, in our model of social learning, an agent located at a vertex v decides upon the technology to adopt for a particular epoch entirely on the basis of the technologies chosen, and the outcomes achieved thereby, by the agents located at the children of v during the previous epoch. If we identify an agent with the vertex at which it is stationed, then the above implies that a vertex 'learns' from its children. This aspect of our model is inspired by the notion of *pure altruism*, commonly used in models of *intergenerational altruism*, such as in [12] and [70]. As described in [42], which proposes an axiomatic theory of *pure, direct altruism* with *altruistic, forward-looking preferences*, a generation exhibits pure altruism if it derives utility, besides possibly from its own consumptions, from its descendants' overall utility or well-being. Altruism is said to be direct if it directly depends on the well-being of all future generations (for instance, environmental policies chosen with a regard to all future generations reflect direct altruism towards these generations). The approach adopted in [12] is also examined in [74]—this approach to intergenerational altruism assumes that successive generations are linked by recursive altruistic preferences, and that individuals have an altruistic concern *only* for their children, who, in turn, have altruistic

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

feelings towards their own children, and so on. In [13], too, the model studied assumes the dependence of a parent's utility on the children's.

Models for social learning serve as frameworks for the dissemination of technologies throughout a population of agents forming a specified network. In [25], in a population of myopic, memoryless agents stationed at the integer points of $\mathbb{Z}$, each agent, at each time step, performs an experiment using one of two technologies available to them, then observes the technology choices and the corresponding outcomes of its two nearest neighbours and of itself. Two learning rules are considered, with the first allowing an agent to change its technology only if it has experienced a failure, and the second allowing each agent to update its technology on the basis of the neighbourhood average. In [9], each $x \in \mathbb{Z}$ is occupied by an agent, and each agent, at each time step, is allowed to adopt one of two available technologies: B and R. An agent retains the technology it adopted in the previous time step if it achieves success using that technology. Otherwise, it observes the technologies adopted, and the outcomes achieved as a consequence, by its two nearest neighbours, and it opts for a change, say from B to R, only if, among its neighbours and itself, the proportion of successes with R is strictly bigger than the proportion of successes with B. In [102], models for (i) the spread of a contagion, (ii) the spread of social influence, and (iii) social learning are formulated at a high level of generality, allowing there to be essentially any distribution of heterogeneous characteristics among the agents concerned, and key differences in their dynamical structures and in their patterns of acceleration are pointed out. In [62], agents decide whether to adopt a new technology, with unknown pay-offs, on the basis of their prior beliefs and the experiences of their neighbours in the network, and it is shown, with use of the mean-field approximations, that this process of diffusion always has at least one stable equilibrium. In [89], each agent (i) possesses unobservable perspectives, (ii) receives private information about the current state and forms an opinion, and (iii) chooses a target agent (different from itself) and observes the target's opinion.

In [24], an individual outside the network, termed the *firm*, has private information regarding the quality of the technology it seeks to diffuse throughout the network, and may pay some agent in the network, known as an *implant*, to propagate its product or idea. Each agent observes the actions of its neighbours over time in order to decide whether to adopt the new technology. Agents are either *innovators* who *always* adopt the new technology or *standard players* who are fully rational and make decisions based on utility maximization. What the firm knows about the interactions of agents with other agents is restricted to the neighbours of the implant chosen, if any exists. In such a model, learning occurs via strategic choices by agents, and this plays a role in determining whether diffusion occurs throughout the whole population or dies out within some finite distance of the origin. In [17], a game of *innovation diffusion* with multiple competing innovations, such as when multiple companies market competing products using viral marketing, is studied, wherein an individual's decision whether to adopt a product or innovation is highly dependent on the choices made by the individual's peers or neighbours in the underlying social network. In [35], it is proposed that, instead of simply comparing the cost of marketing to a customer with the expected profit to be generated from that customer, one should take into account the customer's *network value*, i.e. the expected profit from sales to customers whose choices the customer may influence directly or in a recursive manner. In [60], the *target set selection problem* is studied for the *decreasing cascade model*, i.e. [60] investigated the choice of the initial set of *active* vertices such that the *cascade* (which is the extent to which an innovation or behaviour spreads through the social network under consideration when the influence that an individual's friends and acquaintances have over that individual's decision to purchase or adopt the innovation is leveraged) resulting

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

from this initial active set is as large as possible in expectation. We also mention here [79], in which a conjecture of [60] is established. In [100], a sparse random network of interacting agents, in which each agent's decision is determined by the actions of its neighbours via a simple threshold rule, is studied, and two regimes, in which the network is susceptible to very large cascades or *global cascades*, are identified. In [64] (see also [63]), a model of diffusion, with two available strategies, is studied wherein each agent, at each point in time, chooses the strategy with the best pay-off, given the strategies currently adopted by its neighbours. In [64], criteria under which one of the strategies becomes widespread throughout a random network are proposed, the game's equilibria are analysed, and conditions under which both strategies may coexist in large connected sets are identified. Numerous other studies exist in the literature on models studying the diffusion of technologies, such as [95], [82], [96], and [81].

A third motivation for studying our model arises from its interpretation as a PTA. PTAs admit a very general definition, such as discussed in [38], but in this case it suffices to define a PTA (on $\mathbb{T}_m$) as a *state machine* that consists of two components, namely (i) a finite set $\mathcal{A}$ of symbols or colours known as the *alphabet* and (ii) a stochastic update rule given by a stochastic matrix $\varphi : \mathcal{A}^m \times \mathcal{A} \to [0, 1]$. For any vertex $v \in V(\mathbb{T}_m)$, letting $v_1, v_2, \ldots, v_m$ denote its children, and conditioned on $(\eta_t(v_1), \eta_t(v_2), \ldots, \eta_t(v_m))$, where $\eta_t(v_i) \in \mathcal{A}$ denotes the symbol assigned to v_i at time step t for each $i \in \{1, 2, \ldots, m\}$, the symbol assigned to v at time step $t + 1$ is a random variable $\eta_{t+1}(v)$ whose probability distribution is given by

$$\mathbb{P}\left[\eta_{t+1}(v) = \alpha \middle| \eta_t(v_1), \eta_t(v_2), \ldots, \eta_t(v_m)\right] = \varphi((\eta_t(v_1), \eta_t(v_2), \ldots, \eta_t(v_m)), \alpha) \quad \text{for each } \alpha \in \mathcal{A}.$$

The update from $\eta_t(v)$ to $\eta_{t+1}(v)$ is assumed to happen independently over all $v \in V(\mathbb{T}_m)$. One of the primary questions asked about a PTA is the enumeration and characterization of all its *fixed points*: a joint law ν of $\{\eta_0(v) : v \in V(\mathbb{T}_m)\}$ is called a *fixed point* of the PTA under consideration if the joint law of $\{\eta_1(v) : v \in V(\mathbb{T}_m)\}$ is also ν. This is also the question we seek to answer in this paper for our model. More on PTAs (and their deterministic counterparts, endowed with update rules that are deterministic), as well as their applications, can be found in, for example, [27], [15], [22], and [28].

We conclude this section with [84], which provides an extensive survey of the tools and techniques—primarily (i) martingales and the stochastic approximation and (ii) dynamical systems and their stochastic counterparts—that are used in the analysis of stochastic processes with reinforcements (such as *urn processes*, *reinforced random walks*, *reinforced diffusions*, and processes with negative reinforcement such as *self-avoiding walks*). Stochastic models involving *learning*, such as the one studied in this paper, often exhibit *reinforced behaviour* (in the sense that the agents are often more inclined to select technologies that seem to bring greater rewards), and we speculate that the array of tools and techniques described in [84] would prove to be very useful in the analysis of many such models that are dictated by complicated learning rules.

1.5. Organization of this paper

The rest of this paper is organized as follows. The proof of Theorem 1 is presented in Section 2, with Subsections 2.1, 2.2, and 2.3 respectively addressing parts 1, 1, and 1. The proof of Theorem 2 is outlined in Section 3. Theorem 3 is proved in Section 4, and this section includes, as part of the proof, the statements and proof of the crucial Theorem 5, Lemma 1, and Theorem 6. In Subsection 4.1, the value of $p(m)$, as defined in the statement of Theorem 3, is explicitly computed, by analytical means, for $m = 3$ and $m = 4$. The proof of Theorem 4

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

is provided in Section 5. Finally, since we are not able to provide a mathematically rigorous analysis even for relatively small values of m (such as $m = 3$) when completely arbitrary values of $(p_B, p_R) \in [0, 1)^2$ are considered, Section 6 is dedicated to a detailed discussion, culminating in a conjecture and supported by several examples, of what we believe happens when we consider general values of p_B and p_R for $m = 3$.

2. Proof of Theorem 1

2.1. Proof of 1 of Theorem 1

Recall, from Subsection 1.2, the description of the model we work with when we consider the absolute majority policy. From (2), we already know that the random variables $C_0(v)$, for $v \in V(\mathbb{T}_m)$, are i.i.d., proving the base case of the inductive argument we plan to outline for the proof of 1, the first assertion of Theorem 1. Let us assume that the random variables $C_t(v)$, for $v \in V(\mathbb{T}_m)$, have been shown to be i.i.d., with their common distribution given by (3), for each $t \leqslant n$, for some $n \in \mathbb{N}$. We now consider $C_{n+1}(v)$, for $v \in V(\mathbb{T}_m)$.

Let $S \subset V(\mathbb{T}_m)$ denote any *finite* subset of vertices of our tree $\mathbb{T}_m$. For each $v \in V(\mathbb{T}_m)$, let $\Gamma(v)$ indicate the set of all children of v. Recall, from Subsection 1.2, the random variables $X_t(v)$ and $Y_t(v)$, for each $v \in V(\mathbb{T}_m)$ and each $t \in \mathbb{N}_0$. Since our induction hypothesis states that the collection $\{C_t(v) : v \in V(\mathbb{T}_m)\}$ consists of i.i.d. random variables for *each* $t \leqslant n$, and since $\Gamma(v) \cap \Gamma(v') = \emptyset$ for all $v, v' \in V(\mathbb{T}_m)$ with $v \neq v'$, the random tuples $(C_n(u) : u \in \Gamma(v))$ are also i.i.d. over *all* $v \in S$. Next, recall from Subsection 1.2 that, for each $t \in \mathbb{N}_0$, the collection $\{X_t(v) : v \in V(\mathbb{T}_m)\} \cup \{Y_t(v) : v \in V(\mathbb{T}_m)\}$ is independent; therefore, the entire collection $\{(X_n(u) : u \in \Gamma(v)) \ : v \in S\} \cup \{Y_n(v) : v \in S\}$ is independent.

From (1), it is evident that $C_{n+1}(v)$ is a function of $C_n(u)$ and $X_n(u)$ for each $u \in \Gamma(v)$, as well as of $Y_n(v)$. The discussion in the previous paragraph implies the independence of $\{(C_n(u) : u \in \Gamma(v)) \, , (X_n(u) : u \in \Gamma(v)) \, , Y_n(v)\}$ over all $v \in S$. This, in turn, implies the independence of $C_{n+1}(v)$ over all $v \in S$. That $C_{n+1}(v)$ has the same distribution for each $v \in V(\mathbb{T}_m)$ becomes evident from (1) (which ensures that the conditional distribution of $C_{n+1}(v)$, conditioned on $C_n(u)$, $u \in \Gamma(v)$, is the same for all $v \in V(\mathbb{T}_m)$) and our induction hypothesis that $\{C_n(v) : v \in V(\mathbb{T}_m)\}$ forms an i.i.d. collection.

2.2. Proof of 1 of Theorem 1

Our having established 1 of Theorem 1, let the distribution of $C_t(v)$ be as described by (3), for each $v \in V(\mathbb{T}_m)$ and each $t \in \mathbb{N}$. Recall that ν_t is the joint law of $\{C_t(v) : v \in V(\mathbb{T}_m)\}$.

We first show that if the sequence $\{\pi_t\}_{t \in \mathbb{N}_0}$ converges to a limit π then $\{\nu_t\}_{t \in \mathbb{N}_0}$ converges to the law ν of the i.i.d. collection of random variables $\{C_\infty(v) : v \in V(\mathbb{T}_m)\}$, where the common distribution of each $C_\infty(v)$ is as given in (4). Note that each ν_t is defined on $\{B, R\}^{V(\mathbb{T}_m)}$, which is compact by Tychonoff's theorem (since the set $\{B, R\}$ is compact), so the sequence $\{\nu_t\}_{t \in \mathbb{N}_0}$ is already tight. Therefore, to show that $\{\nu_t\}_{t \in \mathbb{N}_0}$ converges to ν, it suffices to show that all finite-dimensional marginals of $\{\nu_t\}_{t \in \mathbb{N}_0}$ converge to the corresponding marginals of ν.

To this end, for any $k \in \mathbb{N}$, let $v_1, v_2, \ldots, v_k$ be *any* k vertices in $V(\mathbb{T}_m)$. Under ν_t, the joint law of $(C_t(v_1), C_t(v_2), \ldots, C_t(v_k))$ is given by

$$\nu_t\Big|_{v_1, v_2, \ldots, v_k}(\eta_1, \eta_2, \ldots, \eta_k) = \pi_t^{\sum_{i=1}^k \mathbf{1}_{\eta_i = B}} (1 - \pi_t)^{k - \sum_{i=1}^k \mathbf{1}_{\eta_i = B}} \, , \tag{7}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

for each $(\eta_1, \eta_2, \ldots, \eta_k) \in \{B, R\}^k$. Since $\pi_t \to \pi$ as $t \to \infty$, and the function on the right-hand side of (7) is a continuous one (it is a polynomial in π_t),

$$\lim_{t\to\infty} \pi_t^{\sum_{i=1}^k \mathbf{1}_{\eta_i=B}} (1-\pi_t)^{k-\sum_{i=1}^k \mathbf{1}_{\eta_i=B}} = \pi^{\sum_{i=1}^k \mathbf{1}_{\eta_i=B}} (1-\pi)^{k-\sum_{i=1}^k \mathbf{1}_{\eta_i=B}}$$
$$= \nu\Big|_{\nu_1,\nu_2,\ldots,\nu_k}(\eta_1, \eta_2, \ldots, \eta_k),$$

for each $(\eta_1, \eta_2, \ldots, \eta_k) \in \{B, R\}^k$, thus completing the proof of one side of the two-way implication in 1.

Conversely, let $\{\nu_t\}_{t\in\mathbb{N}_0}$ converge to a distributional limit ν as $t \to \infty$. Fix *any* vertex v of $\mathbb{T}_m$, and let $C_\infty(v)$ denote the random variable that indicates the state of v under the measure ν. Since $C_\infty(v)$ takes only the values B and R, its distribution must be of the form given in (4), i.e. $C_\infty(v) = B$ with probability π and $C_\infty(v) = R$ with probability $1 - \pi$. Note that the probability distribution of $C_\infty(v)$ is the marginal of the law ν when we restrict ourselves to the vertex v, just as the probability distribution of $C_t(v)$ is the marginal of ν_t when we restrict ourselves to v. Since ν_t converges to ν as $t \to \infty$, we conclude that $C_t(v)$ converges in distribution to $C_\infty(v)$ as $t \to \infty$. Since $C_t(v)$ is of the form given by (3), we conclude that $\pi_t \to \pi$ as $t \to \infty$.

2.3. Proof of 1 of Theorem 1

We begin by establishing a recurrence relation connecting π_t with π_{t+1}, for each $t \in \mathbb{N}_0$. Fix $v \in V(\mathbb{T}_m)$ and let $v_1, v_2, \ldots, v_m$ denote its children. Fix any $k \in \{0, 1, \ldots, m\}$. From (1), we know that $C_t(v_1), C_t(v_2), \ldots, C_t(v_m)$ are i.i.d., with each of them following the distribution given by 3. Hence, the probability that precisely k of these m random variables equal B and the remaining $m - k$ equal R is given by

$$\mathbb{P}\left[\sum_{i=1}^m \mathbf{1}_{C_t(v_i)=B} = k\right] = \binom{m}{k} \pi_t^k (1-\pi_t)^{m-k}. \tag{8}$$

Conditioned on the event that $C_t(v_i) = B$ for each $i \in \{1, 2, \ldots, k\}$ and that $C_t(v_j) = R$ for each $j \in \{k+1, \ldots, m\}$, the random variables $X_t(v_1), \ldots, X_t(v_k)$ are i.i.d. Bernoulli(p_B), while the random variables $X_t(v_{k+1}), \ldots, X_t(v_m)$ are i.i.d. Bernoulli(p_R). The absolute majority policy, defined in (1), yields

$$\begin{aligned}
&\mathbb{P}\left[C_{t+1}(v) = B \middle| \text{precisely } k \text{ children of } v \text{ are in state } B \text{ at time step } t\right] \\
&= \mathbb{P}\left[C_{t+1}(v) = B \middle| \sum_{i=1}^m \mathbf{1}_{C_t(v_i)=B} = k\right] \\
&= \mathbb{P}\left[C_{t+1}(v) = B \middle| C_t(v_i) = B \text{ for } i \in \{1, \ldots, k\},\ C_t(v_j) = R \text{ for } j \in \{k+1, \ldots, m\}\right] \\
&= \mathbb{P}\left[\sum_{i=1}^k X_t(v_i) > \sum_{j=k+1}^m X_t(v_j)\right] + \mathbb{P}\left[\sum_{i=1}^k X_t(v_i) = \sum_{j=k+1}^m X_t(v_j),\ Y_t(v) = 1\right] \\
&= \mathbb{P}\left[A_k > B_{m-k}\right] + \frac{1}{2}\,\mathbb{P}\left[A_k = B_{m-k}\right],
\end{aligned} \tag{9}$$

where we set $A_k = \sum_{i=1}^k X_t(v_i)$ and $B_{m-k} = \sum_{j=k+1}^m X_t(v_j)$, so, conditioned on the event mentioned above, A_k follows Binomial(k, p_B) and B_{m-k} follows Binomial($m - k$, p_R). When $k = 0$ (respectively when $k = m$), the random variable A_0 (respectively B_0) is simply degenerate at 0.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

The expression obtained in the last step in (9) is precisely $f_m(k)$ as defined in (6). From (8) and (9), the unconditional probability of v being in state B at time step $t+1$ is given by

$$\begin{aligned}\pi_{t+1} &= \mathbb{P}\left[C_{t+1}(v)=B\right] = \sum_{k=0}^{m} \mathbb{P}\left[C_{t+1}(v)=B \,\middle|\, \sum_{i=1}^{m} \mathbf{1}_{C_t(v_i)=B}=k\right] \mathbb{P}\left[\sum_{i=1}^{m} \mathbf{1}_{C_t(v_i)=B}=k\right] \\ &= \sum_{k=0}^{m} \left\{\mathbb{P}\left[A_k > B_{m-k}\right] + \frac{1}{2}\,\mathbb{P}\left[A_k = B_{m-k}\right]\right\} \binom{m}{k} \pi_t^k (1-\pi_t)^{m-k} \\ &= \sum_{k=0}^{m} f_m(k) \binom{m}{k} \pi_t^k (1-\pi_t)^{m-k},\end{aligned}$$

so, from (5), we can write

$$\pi_{t+1} = g_m(\pi_t) \text{ for each } t \in \mathbb{N}_0. \tag{10}$$

When $\pi = \lim_{t\to\infty} \pi_t$ exists, taking the limit, as $t \to \infty$, on both sides of (10) and using the fact that g_m is a polynomial and hence a continuous function, we obtain

$$\pi = g_m(\pi),$$

i.e. we conclude that π is a fixed point of g_m, in the interval [0, 1]. This concludes the proof of 1 of Theorem 1.

3. Proof of Theorem 2

The crux of the proof lies in studying the function g_2. From (6), we see that

$$\begin{aligned}f_2(0) &= \frac{1}{2}\,\mathbb{P}\left[B_2=0\right] = \frac{1}{2}(1-p_R)^2, \\ f_2(1) &= \mathbb{P}\left[A_1=1, B_1=0\right] + \frac{1}{2}\,\mathbb{P}\left[A_1=B_1=1\right] + \frac{1}{2}\,\mathbb{P}\left[A_1=B_1=0\right] \\ &= p_B(1-p_R) + \frac{p_B p_R}{2} + \frac{(1-p_B)(1-p_R)}{2}, \\ f_2(2) &= \mathbb{P}\left[A_2 \geqslant 1\right] + \frac{1}{2}\,\mathbb{P}\left[A_2=0\right] = 1 - \frac{1}{2}\,\mathbb{P}\left[A_2=0\right] = 1 - \frac{(1-p_B)^2}{2}.\end{aligned}$$

Substituting these in (5), we obtain

$$\begin{aligned}g_2(x) &= \frac{1}{2}(1-p_R)^2(1-x)^2 + 2\left\{p_B(1-p_R) + \frac{p_B p_R}{2} + \frac{(1-p_B)(1-p_R)}{2}\right\} x(1-x) \\ &\quad + \left\{1 - \frac{(1-p_B)^2}{2}\right\} x^2,\end{aligned}$$

so

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} g_2(x) = (1-p_R)^2 - 4p_B(1-p_R) - 2p_B p_R - 2(1-p_B)(1-p_R) + 2 - (1-p_B)^2 = p_R^2 - p_B^2,$$

thus implying that g_2 is strictly convex throughout [0, 1] for $p_R > p_B$, strictly concave throughout [0, 1] for $p_R < p_B$, and linear when $p_B = p_R$. Furthermore, we have

$$g_2(0) = \frac{1}{2}(1-p_R)^2 > 0 \text{ for each } p_R \in [0,1) \text{ and } g_2(1) = 1 - \frac{(1-p_B)^2}{2} < 1 \text{ for all } p_B \in [0,1).$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

Consequently, the curve $y = g_2(x)$ lies *above* the line $y = x$ at $x = 0$ and *beneath* the line $y = x$ at $x = 1$ when $p_B, p_R \in [0, 1)$. The two observations mentioned above, when combined, allow us to conclude that there is a unique point of intersection between $y = g_2(x)$ and $y = x$. When $p_R = 1$ and $p_B \in [0, 1)$, this unique point of intersection is $x = 0$, whereas when $p_B = 1$ and $p_R \in [0, 1)$, it is $x = 1$.

Let us now focus on the special case of $p_B = p_R = p \in [0, 1]$. This yields

$$g_2(x) = \frac{1}{2}(1-p)^2 + (2p - p^2)x,$$

which, in turn, yields the unique fixed point $x = 1/2$ when $p \in [0, 1)$, whereas when $p = 1$, $y = g_2(x)$ coincides with the line $y = x$, and therefore, the entire $[0, 1]$ constitutes the set of fixed points of g_2 in that case.

4. Proof of Theorem 3

We begin by summarizing the broad steps via which we accomplish the proof of Theorem 3 (for each fixed but arbitrary $m \in \mathbb{N}$ with $m \geqslant 2$).

(i) We begin by observing that $1/2$ is always a fixed point of g_m, and that $\alpha \in [0, 1]$ is a fixed point of g_m if and only if $1 - \alpha$ is also a fixed point of g_m.

(ii) In Theorem 5, we establish that the function g_m is strictly convex on $[0, 1/2]$ and strictly concave on $[1/2, 1]$ for every $p \in (0, 1]$ (for $p = 0$, we show that $g_m(x) = 1/2$ for all $x \in [0, 1]$). This, in turn, allows us to conclude that the curve $y = g_m(x)$ may intersect the line $y = x$ at most twice in the interval $[0, 1/2]$, and at most twice in $[1/2, 1]$. Because of 4, we have two possibilities:

  (a) g_m has a unique fixed point, $1/2$, in the interval $[0, 1]$;

  (b) g_m has precisely three fixed points in $[0, 1]$, and these are of the form α, $1/2$, and $1 - \alpha$, for some $\alpha \in [0, 1/2)$.

(iii) We then show, in Lemma 1, that g_m has a unique fixed point in $[0, 1]$ if and only if its slope at $1/2$ is bounded above by 1, i.e. $g'_m(1/2) \leqslant 1$.

(iv) In Theorem 6, we show that the derivative $g'_m(1/2)$ of g_m at $1/2$ is strictly increasing as a function of p, for $p \in (0, 1)$. This is followed by our showing that $g'_m(1/2) = 0$ at $p = 0$, and $g'_m(1/2) > 1$ at $p = 1$. Together, these guarantee the existence of a unique $p(m) \in (0, 1)$ such that $g'_m(1/2) \leqslant 1$ for all $p \in [0, p(m)]$ and $g'_m(1/2) > 1$ for all $p \in (p(m), 1]$.

(v) From the final conclusion of 4, from 4, and from 4, we conclude that:

  (a) g_m has a unique fixed point in $[0, 1]$ for each $p \in [0, p(m)]$;

  (b) g_m has precisely three fixed points, of the form α, $1/2$, and $1 - \alpha$, for some $\alpha = \alpha(p) < 1/2$, for each $p \in (p(m), 1]$.

  This completes the proofs of 3 and 3 of Theorem 3.

As outlined above, we begin with a couple of simple observations, as follows. When $p_B = p_R = p$, in the definition of f_m in (6), the random variable A_k follows Binomial(k, p) and the random variable B_{m-k} follows Binomial$(m - k, p)$, for each $k \in \{0, 1, \ldots, m\}$. This allows us to write

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$f_m(m-k) = \mathbb{P}\left[A_{m-k} > B_k\right] + \frac{1}{2}\mathbb{P}\left[A_{m-k} = B_k\right]$$
$$= 1 - \mathbb{P}\left[A_{m-k} < B_k\right] - \mathbb{P}\left[A_{m-k} = B_k\right] + \frac{1}{2}\mathbb{P}\left[A_{m-k} = B_k\right]$$
$$= 1 - \mathbb{P}\left[A_k > B_{m-k}\right] - \frac{1}{2}\mathbb{P}\left[A_k = B_{m-k}\right] = 1 - f_m(k). \tag{11}$$

In particular, when m is even, this implies that $f_m(m/2) = 1/2$. We refer to this condition, which boils down to the identity $f_m(k) + f_m(m-k) = 1$ for each $k \in \{0, 1, \ldots, m\}$, as the *symmetry criterion*. Next, using (11), we have, for any $\alpha \in [0, 1]$,

$$g(\alpha) = \sum_{k=0}^{m} f_m(k)\binom{m}{k}\alpha^k(1-\alpha)^{m-k} = \sum_{k=0}^{m}\{1 - f_m(m-k)\}\binom{m}{k}\alpha^k(1-\alpha)^{m-k}$$
$$= \sum_{k=0}^{m}\binom{m}{k}\alpha^k(1-\alpha)^{m-k} - \sum_{k=0}^{m} f_m(m-k)\binom{m}{m-k}\alpha^k(1-\alpha)^{m-k} = 1 - g(1-\alpha),$$

which immediately tells us that when $p_B = p_R$:

(i) $1/2$ is *always* a fixed point of g_m;

(ii) $\alpha \in [0, 1]$ is a fixed point of g_m if and only if $1 - \alpha$ is as well.

This completes 4, and we now come to 4. To prove Theorem 3, it is crucial that we understand what the function g_m looks like in the interval $[0, 1]$. This is what is captured in Theorem 5.

Theorem 5. *For each $p \in (0, 1]$, the curve g_m, defined as in (5) (with f_m as defined in (6)), is strictly convex on the interval $[0, 1/2]$ and strictly concave on the interval $[1/2, 1]$.*

Remark 3. Note that, at $p = 0$, $f_m(k) = 1/2$ for each $k \in \{0, 1, \ldots, m\}$, as is evident from (9) (since each of A_k and B_{m-k}, in (9), is degenerate at 0 when $p = 0$). Thus, (5) yields $g_m(x) = 1/2$, a constant function. Therefore, the conclusion of Theorem 5 trivially holds for $p = 0$. Moreover, this shows that g_m has only one fixed point in the interval $[0, 1]$, namely $1/2$, when $p = 0$.

Proof. The crux of the proof lies in determining the sign of $g''(x)$ for $x \in [0, 1/2]$ and for $x \in [1/2, 1]$, separately. To this end, we note, from (5), that

$$g_m'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{k=0}^{m} f_m(k)\binom{m}{k}x^k(1-x)^{m-k} = m\sum_{\ell=0}^{m-1}\{f_m(\ell+1) - f_m(\ell)\}\binom{m-1}{\ell}x^\ell(1-x)^{m-1-\ell}. \tag{12}$$

Iterating the same argument that leads to (12), we obtain

$$g_m''(x) = m(m-1)\sum_{\ell=0}^{m-2}\{f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)\}\binom{m-2}{\ell}x^\ell(1-x)^{m-2-\ell}. \tag{13}$$

So far we have not used any property specific to the function f_m, i.e. any property pertaining to the absolute majority policy. In what follows, we use (11). The notation $\lceil x \rceil$, for any $x \in \mathbb{R}$,

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

indicates the smallest integer that is greater than or equal to the real number x, and the notation $\lfloor x \rfloor$ indicates the largest integer that is bounded above by x. When m is even, we note that (11) (along with the identity $f_m(m/2) = 1/2$ that was stated immediately after (11)) yields

$$f_m\left(\frac{m}{2}+1\right) - 2f_m\left(\frac{m}{2}\right) + f_m\left(\frac{m}{2}-1\right) = 1 - 2 \cdot \frac{1}{2} = 0,$$

so the summand corresponding to $\ell = \lceil \frac{m}{2} \rceil - 1$ in (13) equals 0 when m is even. This observation, along with the application of (11), allows us to conclude the following:

$$\begin{aligned} g''(x) &= m(m-1) \sum_{\ell=0}^{\lceil \frac{m}{2} \rceil - 2} \{f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)\} \binom{m-2}{\ell} x^\ell (1-x)^{m-2-\ell} \\ &\quad + \sum_{\ell=\lceil \frac{m}{2} \rceil - 1}^{m-2} \{f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)\} \binom{m-2}{\ell} x^\ell (1-x)^{m-2-\ell} \\ &= m(m-1) \sum_{\ell=0}^{\lceil \frac{m}{2} \rceil - 2} \{f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)\} \binom{m-2}{\ell} x^\ell (1-x)^{\ell} \\ &\quad \left\{(1-x)^{m-2-2\ell} - x^{m-2-2\ell}\right\}. \end{aligned}$$

The difference $(1-x)^{m-2-2\ell} - x^{m-2-2\ell}$ is strictly positive for $x \in [0, 1/2)$, and strictly negative for $x \in (1/2, 1]$, for each $\ell \in \left\{0, 1, \ldots, \lceil \frac{m}{2} \rceil - 2\right\}$ (this range of values for ℓ ensures that the exponent $(m-2-2\ell)$ is strictly positive). To prove Theorem 5, it thus suffices to show that

$$f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell) > 0 \quad \text{for each } \ell \in \left\{0, 1, \ldots, \left\lceil \frac{m}{2} \right\rceil - 2\right\} \text{ and each } p \in (0, 1]. \tag{14}$$

The rest of the proof of Theorem 5 is dedicated to establishing (14). We first focus on $p \in (0, 1)$, and prove (14) for the special case of $p = 1$ later.

To begin with, we express the left side of the inequality in (14) in terms of $X_t(v_1), X_t(v_2), \ldots, X_t(v_m)$ (where $v_1, v_2, \ldots, v_m$ are the children of an arbitrary but fixed vertex v of $\mathbb{T}_m$), via (9). We consider the event in which $C_t(v_i) = B$ for each $i \in \{1, 2, \ldots, \ell\}$ and $C_t(v_i) = R$ for each $i \in \{\ell+3, \ldots, m\}$ (similar to the event conditioned on before we began the derivation of (9)). The idea we use is as follows.

(i) When computing $f_m(\ell)$, we assume that $C_t(v_{\ell+1}) = C_t(v_{\ell+2}) = R$; when computing $f_m(\ell+1)$, we assume that $C_t(v_{\ell+1}) = B$ and $C_t(v_{\ell+2}) = R$; when computing $f_m(\ell+2)$, we assume that $C_t(v_{\ell+1}) = C_t(v_{\ell+2}) = B$.

(ii) We then consider various possible scenarios (in terms of the values of $\sum_{i=1}^{\ell} X_t(v_i)$ and $\sum_{i=\ell+3}^{m} X_t(v_i)$, relative to each other), and determine the extent to which each such scenario contributes to each of the conditional probabilities $f_m(\ell)$, $f_m(\ell+1)$, and $f_m(\ell+2)$.

(iii) Finally, we compute the contribution of each such scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$, and add everything to obtain our final expression.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

Scenario 1: When $\sum_{i=1}^{\ell} X_t(v_i) > \sum_{i=\ell+3}^{m} X_t(v_i) + 2$, we have $\sum_{i=\ell+1}^{m} X_t(v_i) < \sum_{i=1}^{\ell} X_t(v_i)$ no matter what the values of $X_t(v_{\ell+1})$ and $X_t(v_{\ell+2})$ are, which further implies $\sum_{i=\ell+2}^{m} X_t(v_i) < \sum_{i=1}^{\ell+1} X_t(v_i)$ and $\sum_{i=\ell+3}^{m} X_t(v_i) < \sum_{i=1}^{\ell+2} X_t(v_i)$. Therefore, the contribution of this scenario to each of $f_m(\ell)$, $f_m(\ell+1)$, and $f_m(\ell+2)$ is $\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) > \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right]$. Consequently, the contribution of this scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is 0.

Scenario 2: When $\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2$, the strict inequality $\sum_{i=\ell+1}^{m} X_t(v_i) < \sum_{i=1}^{\ell} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0), (0,1), (1,0)\}$, and it becomes an equality if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)$. Thus, the contribution of this scenario to $f_m(\ell)$, from (9), is

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0), (0,1), (1,0)\}\right]$$
$$+ \frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right].$$

On the other hand, we have $\sum_{i=\ell+2}^{m} X_t(v_i) < \sum_{i=1}^{\ell+1} X_t(v_i)$ no matter what the values of $X_t(v_{\ell+1})$ and $X_t(v_{\ell+2})$ are, further implying that $\sum_{i=\ell+3}^{m} X_t(v_i) < \sum_{i=1}^{\ell+2} X_t(v_i)$. Thus, the contribution of this scenario to each of $f_m(\ell+1)$ and $f_m(\ell+2)$ is $\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right]$. The net contribution of this scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is

$$\begin{aligned}
&\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right] - 2\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right] \\
&\quad + \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0), (0,1), (1,0)\}\right] \\
&\quad + \frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right] \\
&= -\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right] \\
&= -\frac{p^2}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right].
\end{aligned}$$

Scenario 3: When $\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1$, the strict inequality $\sum_{i=\ell+1}^{m} X_t(v_i) < \sum_{i=1}^{\ell} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)$, and it becomes an equality if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1), (1,0)\}$. Thus, the contribution of this scenario to $f_m(\ell)$ is

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0)\}\right].$$

Next, the strict inequality $\sum_{i=\ell+2}^{m} X_t(v_i) < \sum_{i=1}^{\ell+1} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0),(1,0),(1,1)\}$, and it becomes an equality if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,1)$. Thus, the contribution of this scenario to $f_m(\ell+1)$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0),(1,0),(1,1)\}\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,1)\right].$$

Finally, since $\sum_{i=\ell+3}^{m} X_t(v_i) < \sum_{i=1}^{\ell+2} X_t(v_i)$ no matter what values $X_t(v_{\ell+1})$ and $X_t(v_{\ell+2})$ assume, the contribution of this scenario to $f_m(\ell+2)$ is $\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1\right]$. The net contribution of this scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0)\}\right]$$
$$-2\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0),(1,0),(1,1)\}\right]$$
$$-\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,1)\right]$$
$$+\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1\right]$$
$$= -\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,1)\right]$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$- \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right]$$

$$= -p^2\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1\right].$$

Scenario 4: When $\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i)$, the strict inequality $\sum_{i=\ell+1}^{m} X_t(v_i) < \sum_{i=1}^{\ell} X_t(v_i)$ never holds, but an equality does if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)$. The contribution of this scenario to $f_m(\ell)$ is thus

$$\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right].$$

Next, the strict inequality $\sum_{i=\ell+2}^{m} X_t(v_i) < \sum_{i=1}^{\ell+1} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}, X_t(v_{\ell+2})) = (1,0)$, and it becomes an equality if and only if $(X_t(v_{\ell+1}, X_t(v_{\ell+2})) \in \{(0,0),(1,1)\}$. The contribution of this scenario to $f_m(\ell+1)$ is thus

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right]$$

$$+ \frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0),(1,1)\}\right].$$

Finally, the strict inequality $\sum_{i=\ell+3}^{m} X_t(v_i) < \sum_{i=1}^{\ell+2} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}, X_t(v_{\ell+2})) \in \{(0,1),(1,0),(1,1)\}$, and it becomes an equality if and only if $(X_t(v_{\ell+1}, X_t(v_{\ell+2})) = (0,0)$. Therefore, the contribution of this scenario to $f_m(\ell+2)$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0),(1,1)\}\right]$$

$$+ \frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right].$$

The net contribution of this scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is thus

$$\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right]$$

$$- 2\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right]$$

$$- \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,0),(1,1)\}\right]$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$+\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0),(1,1)\}\right]$$

$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,0)\right]$$

$$= -\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right]$$

$$+\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i),\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (0,1)\right] = 0.$$

Scenario 5: When $\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1$, we have $\sum_{i=1}^{\ell} X_t(v_i) < \sum_{i=\ell+1}^{m} X_t(v_i)$ no matter what the values of $X_t(v_{\ell+1})$ and $X_t(v_{\ell+2})$ are, so this scenario contributes nothing to $f_m(\ell)$. Next, we have $\sum_{i=\ell+2}^{m} X_t(v_i) = \sum_{i=1}^{\ell+1} X_t(v_i)$ if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)$, and under no circumstances do we have the strict inequality $\sum_{i=\ell+2}^{m} X_t(v_i) < \sum_{i=1}^{\ell+1} X_t(v_i)$. Thus, the contribution of this scenario to $f_m(\ell+1)$ is

$$\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right].$$

Finally, the strict inequality $\sum_{i=\ell+3}^{m} X_t(v_i) < \sum_{i=1}^{\ell+2} X_t(v_i)$ holds if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)$, and this becomes an equality if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0)\}$. Thus, the contribution of this scenario to $f_m(\ell+2)$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right]$$

$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0)\}\right].$$

The net contribution of this scenario to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is thus

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right]$$

$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) \in \{(0,1),(1,0)\}\right]$$

$$-\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,0)\right]$$

$$= p^2\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1\right].$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

Scenario 6: When $\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2$, we have $\sum_{i=\ell+2}^{m} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + X_t(v_{\ell+2}) = \sum_{i=1}^{\ell} X_t(v_i) + 2 + X_t(v_{\ell+2}) > \sum_{i=1}^{\ell} X_t(v_i) + X_t(v_{\ell+1}) = \sum_{i=1}^{\ell+1} X_t(v_i)$ no matter what the values of $X_t(v_{\ell+1})$ and $X_t(v_{\ell+2})$ are. This inequality further implies that $\sum_{i=\ell+1}^{m} X_t(v_i) > \sum_{i=1}^{\ell} X_t(v_i)$. Thus, this scenario has no contribution to $f_m(\ell+1)$ or to $f_m(\ell)$. We have $\sum_{i=\ell+3}^{m} X_t(v_i) = \sum_{i=1}^{\ell+2} X_t(v_i)$ if and only if $(X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1, 1)$. Thus, the contribution of this scenario to $f_m(\ell+2)$, and to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$, is

$$\begin{aligned}&\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2,\ (X_t(v_{\ell+1}), X_t(v_{\ell+2})) = (1,1)\right]\\ &\quad = \frac{p^2}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right].\end{aligned}$$

Scenario 7: When $\sum_{i=1}^{\ell} X_t(v_i) < \sum_{i=\ell+3}^{m} X_t(v_i) - 2$, it is straightforward to see that it contributes to none of $f_m(\ell)$, $f_m(\ell+1)$, and $f_m(\ell+2)$, so its net contribution to $f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)$ is 0.

Combining the contributions obtained from all the scenarios considered above, we conclude that

$$\begin{aligned}&f_m(\ell+2) - 2f_m(\ell+1) + f_m(\ell)\\ &\quad = p^2\left\{\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1\right] - \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1\right]\right\}\\ &\qquad + \frac{p^2}{2}\left\{\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right] - \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right]\right\}. \quad (15)\end{aligned}$$

Since $\sum_{i=1}^{\ell} X_t(v_i)$ follows Binomial(ℓ, p), $\sum_{i=\ell+3}^{m} X_t(v_i)$ follows Binomial$(m-\ell-2, p)$, and they are independent of each other, we have (the strict inequality holds because $p \in (0,1)$, making the omitted term strictly positive)

$$\begin{aligned}&\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 1\right] - \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 1\right]\\ &= \sum_{r=0}^{\ell} \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = r,\ \sum_{i=\ell+3}^{m} X_t(v_i) = r+1\right] - \sum_{r=0}^{\ell-1} \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = r+1,\ \sum_{i=\ell+3}^{m} X_t(v_i) = r\right]\\ &> \sum_{r=0}^{\ell-1} \binom{\ell}{r} p^r (1-p)^{\ell-r} \binom{m-2-\ell}{r+1} p^{r+1} (1-p)^{m-3-\ell-r}\\ &\quad - \sum_{r=0}^{\ell-1} \binom{\ell}{r+1} p^{r+1} (1-p)^{\ell-r-1} \binom{m-2-\ell}{r} p^r (1-p)^{m-2-\ell-r}\end{aligned}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$= \sum_{r=0}^{\ell-1} \frac{\ell!(m-2-\ell)!}{r!(r+1)!(\ell-r)!(m-2-r-\ell)!} \{(m-2-r-\ell)-(\ell-r)\}\, p^{2r+1}(1-p)^{m-3-2r}$$

$$= \sum_{r=0}^{\ell-1} \frac{\ell!(m-2-\ell)!(m-2-2\ell)}{r!(r+1)!(\ell-r)!(m-2-r-\ell)!} p^{2r+1}(1-p)^{m-3-2r}, \tag{16}$$

and this is strictly positive for all $p \in (0,1)$ since $\ell \leqslant \lceil \frac{m}{2} \rceil - 2 \implies m-2-2\ell \geqslant 1$. This takes care of the first difference in the expression in (15).

We deal with the second difference in (15) in much the same manner as the first, but with a couple of additional observations, as follows.

(i) When m is odd and $\ell = \lceil \frac{m}{2} \rceil - 2$, we have $\ell + 2 > m - \ell - 2 = \ell + 1$, which means that the event $\left\{\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right\}$ equals the union $\bigcup_{r=0}^{\ell-1} \left\{\sum_{i=1}^{\ell} X_t(v_i) = r, \sum_{i=\ell+3}^{m} X_t(v_i) = r+2\right\}$, whereas when $\ell < \lceil \frac{m}{2} \rceil - 2$, it equals the union $\bigcup_{r=0}^{\ell} \left\{\sum_{i=1}^{\ell} X_t(v_i) = r, \sum_{i=\ell+3}^{m} X_t(v_i) = r+2\right\}$.

(ii) When m is even, the event $\left\{\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right\}$ equals the union $\bigcup_{r=0}^{\ell} \left\{\sum_{i=1}^{\ell} X_t(v_i) = r, \sum_{i=\ell+3}^{m} X_t(v_i) = r+2\right\}$ for *each* $\ell \in \{0, 1, \ldots, \lceil \frac{m}{2} \rceil - 2\}$.

We can thus write $\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right] \geqslant \sum_{r=0}^{\ell-1} \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = r, \sum_{i=\ell+3}^{m} X_t(v_i) = r+2\right]$. Therefore, the second difference in the expression in (15) becomes (once again, the strict inequality holds because $p \in (0,1)$, making the omitted term strictly positive)

$$\mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) - 2\right] - \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = \sum_{i=\ell+3}^{m} X_t(v_i) + 2\right]$$

$$\geqslant \sum_{r=0}^{\ell-1} \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = r, \sum_{i=\ell+3}^{m} X_t(v_i) = r+2\right] - \sum_{r=0}^{\ell-2} \mathbb{P}\left[\sum_{i=1}^{\ell} X_t(v_i) = r+2, \sum_{i=\ell+3}^{m} X_t(v_i) = r\right]$$

$$> \sum_{r=0}^{\ell-2} \binom{\ell}{r} p^r (1-p)^{\ell-r} \binom{m-\ell-2}{r+2} p^{r+2}(1-p)^{m-\ell-r-4}$$

$$- \sum_{r=0}^{\ell-2} \binom{\ell}{r+2} p^{r+2}(1-p)^{\ell-r-2} \binom{m-\ell-2}{r} p^r (1-p)^{m-\ell-r-2}$$

$$= \sum_{r=0}^{\ell-2} \frac{\ell!(m-2-\ell)!\,\{(m-\ell-r-3)(m-\ell-r-2)-(\ell-r-1)(\ell-r)\}}{r!(r+2)!(\ell-r)!(m-\ell-r-2)!} p^{2r+2}(1-p)^{m-2r-4}, \tag{17}$$

and this is strictly positive for $p \in (0,1)$ because, as noted above, $\ell \leqslant \lceil \frac{m}{2} \rceil - 2 \implies m-2-2\ell \geqslant 1$, which in turn implies that $m-\ell-r-2 > \ell-r$ as well as $m-\ell-r-3 > \ell-r-1$.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

Since we have established that the expression in (16) as well as that in (17) is strictly positive for each $\ell \in \left\{0, 1, \ldots, \left\lceil \frac{m}{2} \right\rceil - 2\right\}$ and for each $p \in (0, 1)$, so is the expression in (15), which in turn accomplishes our objective of proving (14) for $p \in (0, 1)$.

We now come to the proof of (14) for $p = 1$. Note, in this case, from (9), that (i) $f_m(\ell) = 1$ if $\ell > m - \ell$, (ii) $f_m(\ell) = 1/2$ if $\ell = m - \ell$ (which happens only if m is even), and (iii) $f_m(\ell) = 0$ if $\ell < m - \ell$. Note that $\ell \leqslant \left\lceil \frac{m}{2} \right\rceil - 2 \implies \ell + 1 \leqslant m - \ell - 2 < m - (\ell + 1)$, so $f_m(\ell + 1) = 0$, further implying $f_m(\ell) = 0$, for each $\ell \in \left\{0, 1, \ldots, \left\lceil \frac{m}{2} \right\rceil - 2\right\}$. Moreover, $\ell < \left\lceil \frac{m}{2} \right\rceil - 2 \implies \ell \leqslant \left\lceil \frac{m}{2} \right\rceil - 3 \implies \ell + 2 \leqslant m - \ell - 3 < m - (\ell + 2)$, so $f_m(\ell + 2) = 0$ for all $\ell < \left\lceil \frac{m}{2} \right\rceil - 2$. When $m = 2n$ for some $n \in \mathbb{N}$, and $\ell = \left\lceil \frac{m}{2} \right\rceil - 2 = n - 2$, we have $\ell + 2 = n = m - \ell - 2$, implying that $f_m(\ell + 2) = 1/2$ in this case. When $m = 2n + 1$ for some $n \in \mathbb{N}$, and $\ell = \left\lceil \frac{m}{2} \right\rceil - 2 = n - 1$, we have $\ell + 2 = n + 1 > n = m - \ell - 2$, so $f_m(\ell + 2) = 1$ in this case. All these observations together imply that the expression $f_m(\ell + 2) - 2f_m(\ell + 1) + f_m(\ell)$ equals 0 for each $\ell < \left\lceil \frac{m}{2} \right\rceil - 2$, and when $\ell = \left\lceil \frac{m}{2} \right\rceil - 2$, it equals $1/2$ for m even and 1 for m odd, once again establishing (14).

This completes the proof of (14), and consequently Theorem 5, for all $p \in (0, 1]$. □

Before we proceed any further, we explain why we already have the full picture of g_m when $p = 1$. From Theorem 5, we conclude that the curve $y = g_m(x)$ has at most two points of intersection with the line $y = x$ in *each* of the two intervals $[0, 1/2]$ and $[1/2, 1]$. From 1, we know that $1/2$ is already a fixed point, which means that g_m is allowed to have at most one fixed point in the interval $[0, 1/2)$, and at most one fixed point in the interval $(1/2, 1]$. From (9), we note that (i) $f_m(k) = 1$ for each $k \in \{n + 1, \ldots, 2n + 1\}$ when $m = 2n + 1$ and (ii) $f_m(k) = 1$ for each $k \in \{n + 1, \ldots, 2n\}$ and $f_m(n) = 1/2$ when $m = 2n$, when $p = 1$. This tells us that when $p = 1$, we have $g_m(0) = 0$, and this, along with 4, yields *both* 0 and 1 as fixed points of g_m when $p = 1$. Consequently, 0, $1/2$, and 1 are the three fixed points of g_m when $p = 1$.

Recall from Remark 3 that $y = g_m(x)$ coincides with the line $y = 1/2$ when $p = 0$, so g_m has the unique fixed point $1/2$ in $[0, 1]$ when $p = 0$. We, therefore, need focus only on investigating the number of fixed points of g_m when $p \in (0, 1)$. The following result is the first step towards such an investigation.

Lemma 1. *For $p \in (0, 1)$, the function g_m has a unique fixed point in $[0, 1]$ if and only if its derivative $g_m'(1/2)$ at $1/2$ is bounded above by 1.*

Proof. Note that, from (5) and (9), we have $g_m(0) = f_m(0) = \frac{1}{2}\,\mathbb{P}\,[B_m = 0] = \frac{1}{2}(1 - p)^m > 0$ (since $p < 1$), so the curve $y = g_m(x)$ lies *above* the line $y = x$ at $x = 0$. If α is the smallest fixed point of g_m in the interval $[0, 1]$ (and $\alpha > 0$ since we have just shown that $g_m(0) > 0$), the curve $y = g_m(x)$ travels from *above* $y = x$ to *beneath* $y = x$ at $x = \alpha$. Therefore, the slope $g_m'(\alpha)$ of the curve $y = g_m(x)$ at $x = \alpha$ must be bounded above by the slope of the line $y = x$, which is 1. We already know, from 4, that $1/2$ is a fixed point of g_m. When g_m has a unique fixed point in $[0, 1]$, it must be $\alpha = 1/2$, allowing us to conclude that $g_m'(1/2) \leqslant 1$.

When g_m has multiple fixed points in $[0, 1]$, we know, from 4 and 4, that $\alpha < 1/2$, and that α, $1/2$, and $1 - \alpha$ are all fixed points of g_m. From Theorem 5, we know that g_m has at most two fixed points in $[0, 1/2]$ and at most two in $[1/2, 1]$, so α, $1/2$, and $1 - \alpha$ constitute *all* of the fixed points of g_m in $[0, 1]$. Evidently, the curve $y = g_m(x)$ travels from *above* $y = x$ to *beneath* $y = x$ at $x = \alpha$, whereas it travels from *beneath* $y = x$ to *above* $y = x$ at $x = 1/2$. Consequently, the slope $g_m'(1/2)$ of g_m at $1/2$ must be strictly greater than the slope of the line $y = x$, which is 1, yielding $g_m'(1/2) > 1$ in this case, as desired. This completes the proof. □

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

We now come to the final step, i.e. 4, of proving Theorem 3. From Lemma 1 as well as the second paragraph of its proof, it suffices for us to show that, for each $m \in \mathbb{N}$ with $m \geqslant 2$, there exists $p(m) \in (0, 1)$ such that $g'_m(1/2) \leqslant 1$ for each $p \leqslant p(m)$, and $g'_m(1/2) > 1$ for each $p > p(m)$. This is precisely what is accomplished by proving the following result.

Theorem 6. *For each $m \in \mathbb{N}$ with $m \geqslant 2$, the derivative $g'_m(1/2)$ is a strictly increasing function of the parameter p, for $p \in (0, 1)$.*

Proof. We have, so far, used the notation $f_m(\ell)$, for each $\ell \in \{0, 1, \ldots, m-1\}$, which does *not* make explicit the dependence that these functions (recall their definitions from (6)) have on p. Since this proof constitutes investigating the behaviour of these functions with respect to p, we tweak the notation a little (just for the sake of this proof and not beyond that), and replace $f_m(\ell)$ by $f_m^\ell(p)$, for each $\ell \in \{0, 1, \ldots, m\}$. Likewise, to emphasize the dependence of $g'_m(1/2)$ on the value of p under consideration, we write $g'_m(1/2)\big|_p$ throughout the proof of Theorem 6.

We use (12) and (11) (along with the fact that $f_m^{m/2}(p) = 1/2$ when m is even, stated immediately after (11)) to obtain, for each $p \in (0, 1)$,

$$\begin{aligned} g'_m\left(\frac{1}{2}\right)\Big|_p &= \frac{m}{2^{m-1}} \sum_{\ell=0}^{m-1} \left\{f_m^{\ell+1}(p) - f_m^\ell(p)\right\} \binom{m-1}{\ell} \\ &= \frac{m}{2^{m-2}} \left[\sum_{\ell=0}^{\left\lfloor \frac{m-1}{2} \right\rfloor} \frac{(m-1)!(2\ell - m) f_m^\ell(p)}{\ell!(m-\ell)!} + \binom{m-1}{\left\lfloor \frac{m-1}{2} \right\rfloor}\right]. \end{aligned} \tag{18}$$

Our task is to establish that the derivative of the expression in (18), with respect to p, is strictly positive for each $p \in (0, 1)$. Since $\ell \leqslant \left\lfloor \frac{m-1}{2} \right\rfloor \implies 2\ell - m < 0$, this amounts to showing that the derivative of $f_m(\ell)$ with respect to p, for each $\ell \in \left\{0, 1, \ldots, \left\lfloor \frac{m-1}{2} \right\rfloor\right\}$, is strictly negative for each $p \in (0, 1)$. To this end, we prove the following claim:

$$\frac{\mathrm{d}}{\mathrm{d}p} f_m^\ell(p) = -\frac{m - 2\ell}{2} \sum_{i=0}^{\ell} \binom{m-\ell}{i} \binom{\ell}{i} p^{2i} (1-p)^{m-1-2i} \text{ for each } \ell \in \left\{0, 1, \ldots, \left\lfloor \frac{m-1}{2} \right\rfloor\right\}. \tag{19}$$

To prove (19), we fix an $\ell \in \left\{0, 1, \ldots, \left\lfloor \frac{m-1}{2} \right\rfloor\right\}$ for the rest of the proof.

The idea used in proving (19) is the same as that inspiring Russo's formula (see, for instance, Subsection 2.4 of [51]). Fix a vertex v of $\mathbb{T}_m$, and let $v_1, \ldots, v_m$ denote its children. To take into account how the rate of change of $f_m^\ell(p)$ with respect to p is impacted by the outcome of the experiment performed by the agent at each of $v_1, \ldots, v_m$ at time step t, let us fix $p_1, p_2, \ldots, p_m$, each in $(0, 1)$, and let $X_t(v_1), \ldots, X_t(v_m)$ be mutually independent, but with $X_t(v_i)$ now following Bernoulli(p_i) for each $i \in \{1, 2, \ldots, m\}$. We let $\overline{p} = (p_1, p_2, \ldots, p_m)$, and introduce the generalized definition

$$\begin{aligned} f_m^\ell(\overline{p}) = \mathbb{P}\big[C_{t+1}(v) = B \big| C_t(v_i) = B \quad &\text{for each } i \in \{1, \ldots, \ell\} \\ \text{and} \quad C_t(v_i) = R \quad &\text{for each } i \in \{\ell+1, \ldots, m\}\big], \end{aligned}$$

where the rule for deciding $C_{t+1}(v)$, given $C_t(v_1), \ldots, C_t(v_m)$, is the same as that described in (1). Note that, in our original set-up, we have $p_1 = p_2 = \cdots = p_m = p$, and $f_m^\ell(\overline{p})$ simply boils down to $f_m^\ell(p)$.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

If we fix any $0 \leqslant \Delta(p_i) \leqslant 1 - p_i$ for each $i \in \{1, 2, \ldots, m\}$, and set $\Delta(\overline{p}) = (\Delta(p_1), \Delta(p_2), \ldots, \Delta(p_m))$, then, by the well-known formula for total derivatives, we obtain

$$f_m^\ell\left(\overline{p} + \Delta(\overline{p})\right) - f_m^\ell(\overline{p}) = \sum_{i=1}^{m} \frac{\partial}{\partial p_i} f_m^\ell(\overline{p}) \cdot \Delta(p_i). \tag{20}$$

Note that in our original set-up, where $p_1 = p_2 = \cdots = p_m = p$, we have, due to symmetry,

$$\begin{aligned} &\frac{\partial}{\partial p_1} f_m^\ell(\overline{p})\Big|_{(p,p,\ldots,p)} = \cdots = \frac{\partial}{\partial p_\ell} f_m^\ell(\overline{p})\Big|_{(p,p,\ldots,p)} \text{ and } \frac{\partial}{\partial p_{\ell+1}} f_m^\ell(\overline{p})\Big|_{(p,p,\ldots,p)} \\ &= \cdots = \frac{\partial}{\partial p_m} f_m^\ell(\overline{p})\Big|_{(p,p,\ldots,p)}, \end{aligned} \tag{21}$$

which is why it suffices for us to find only $\frac{\partial}{\partial p_1} f_m^\ell(\overline{p})$ and $\frac{\partial}{\partial p_{\ell+1}} f_m^\ell(\overline{p})$. Furthermore, when $p_1 = p_2 = \cdots = p_m = p$ and $\Delta(p_1) = \Delta(p_2) = \cdots = \Delta(p_m) = \Delta(p)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}p} f_m^\ell(p) = \lim_{\Delta(p)\to 0} \frac{1}{\Delta(p)} \left\{ f_m^\ell\left(\overline{p} + \Delta(\overline{p})\right)\Big|_{(p+\Delta(p),p+\Delta(p),\ldots,p+\Delta(p))} - f_m^\ell(\overline{p})\Big|_{(p,p,\ldots,p)} \right\}. \tag{22}$$

For each $i \in \{1, 2, \ldots, m\}$, we set $\Delta^{(i)}(\overline{p})$ to be the m-tuple in which the ith coordinate equals $\Delta(p_i)$, and every other coordinate equals 0. Note that computing $\frac{\partial}{\partial p_i} f_m^\ell(\overline{p})$ amounts to finding $f_m^\ell\left(\overline{p} + \Delta^{(i)}(\overline{p})\right) - f_m^\ell(\overline{p})$ for each $i \in \{1, 2, \ldots, m\}$.

Let $U_1, U_2, \ldots, U_m$ denote i.i.d. Uniform[0, 1] random variables. When $X_t(v_i)$ follows Bernoulli(p_i), we can write $X_t(v_i) = \mathbf{1}_{U_i \leqslant p_i}$, for each $i \in \{1, 2, \ldots, m\}$. Consequently, analogous to (9), we obtain

$$f_m^\ell(\overline{p}) = \mathbb{P}\left[\sum_{i=1}^{\ell} \mathbf{1}_{U_i \leqslant p_i} > \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}\right] + \frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell} \mathbf{1}_{U_i \leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}\right].$$

We utilize this to now compute $f_m^\ell\left(\overline{p} + \Delta^{(1)}(\overline{p})\right) - f_m^\ell(\overline{p})$, considering various possible scenarios.

(i) When $\sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} > \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}$, since each of $\mathbf{1}_{U_1 \leqslant p_1}$ and $\mathbf{1}_{U_1 \leqslant p_1 + \Delta(p_1)}$ is non-negative, we have $\sum_{i=1}^{\ell} \mathbf{1}_{U_i \leqslant p_i} > \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}$ as well as $\mathbf{1}_{U_1 \leqslant p_1 + \Delta(p_1)} + \sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} > \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}$, which means that $\mathbb{P}\left[\sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} > \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}\right]$ is present in each of $f_m^\ell\left(\overline{p} + \Delta^{(1)}(\overline{p})\right)$ and $f_m^\ell(\overline{p})$. Consequently, it has no contribution to $f_m^\ell\left(\overline{p} + \Delta^{(1)}(\overline{p})\right) - f_m^\ell(\overline{p})$.

(ii) When $\sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}$, its contribution to $f_m^\ell(\overline{p})$ is given by

$$\mathbb{P}\left[U_1 \leqslant p_1, \sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}\right] + \frac{1}{2}\,\mathbb{P}\left[U_1 > p_1, \sum_{i=2}^{\ell} \mathbf{1}_{U_i \leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i \leqslant p_i}\right],$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

whereas its contribution to $f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)$ is given by

$$\mathbb{P}\left[U_1 \leqslant p_1+\Delta(p_1), \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i}\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[U_1 > p_1+\Delta(p_1), \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i}\right].$$

If we combine these, the contribution of this scenario to $f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)-f_m^\ell(\overline{p})$ becomes

$$\frac{1}{2}\,\mathbb{P}\left[p_1 < U_1 \leqslant p_1+\Delta(p_1), \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i}\right]$$
$$=\frac{\Delta(p_1)}{2}\,\mathbb{P}\left[\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i}\right].$$

(iii) When $\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1$, its contribution to $f_m^\ell(\overline{p})$ is given by

$$\frac{1}{2}\,\mathbb{P}\left[U_1 \leqslant p_1, \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1\right],$$

whereas its contribution to $f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)$ is given by

$$\frac{1}{2}\,\mathbb{P}\left[U_1 \leqslant p_1+\Delta(p_1), \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1\right].$$

Thus, the contribution of this scenario to $f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)-f_m^\ell(\overline{p})$ is

$$\frac{1}{2}\,\mathbb{P}\left[p_1 < U_1 \leqslant p_1+\Delta(p_1), \sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1\right]$$
$$=\frac{\Delta(p_1)}{2}\,\mathbb{P}\left[\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1\right].$$

(iv) Finally, when $\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} < \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1$, its contribution to each of $f_m^\ell(\overline{p})$ and $f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)$ is 0, so its contribution to their difference is also 0.

Combining the contributions from the four scenarios considered above, we obtain

$$f_m^\ell\left(\overline{p}+\Delta^{(1)}(\overline{p})\right)-f_m^\ell(\overline{p}) = \frac{\Delta(p_1)}{2}\left\{\mathbb{P}\left[\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i}\right]\right.$$
$$\left.+\,\mathbb{P}\left[\sum_{i=2}^{\ell} \mathbf{1}_{U_i\leqslant p_i} = \sum_{i=\ell+1}^{m} \mathbf{1}_{U_i\leqslant p_i} - 1\right]\right\}. \tag{23}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

We now compute $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)-f_m^\ell(\overline{p})$ via a case-by-case analysis in the same manner as above.

(i) When $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}>\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1$, we have $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}>\sum_{i=\ell+1}^{m}\mathbf{1}_{U_i\leqslant p_i}$ as well as $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}>\mathbf{1}_{U_{\ell+1}\leqslant p_{\ell+1}}+\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}$. Thus, the contribution of this scenario to each of $f_m^\ell(\overline{p})$ and $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)$ is $\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}>\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1\right]$, and therefore its contribution to their difference is 0.

(ii) When $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1$, its contribution to $f_m^\ell(\overline{p})$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1,\,U_{\ell+1}>p_{\ell+1}\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1,\,U_{\ell+1}\leqslant p_{\ell+1}\right].$$

Likewise, the contribution of this scenario to $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)$ is

$$\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1,\,U_{\ell+1}>p_{\ell+1}+\Delta(p_{\ell+1})\right]$$
$$+\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1,\,U_{\ell+1}\leqslant p_{\ell+1}+\Delta(p_{\ell+1})\right].$$

Consequently, the contribution of this scenario to $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)-f_m^\ell(\overline{p})$ is

$$-\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1,\,p_{\ell+1}<U_{\ell+1}\leqslant p_{\ell+1}+\Delta(p_{\ell+1})\right]$$
$$=-\frac{\Delta(p_{\ell+1})}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1\right].$$

(iii) When $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}$, its contribution to $f_m^\ell(\overline{p})$ is

$$\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i},\,U_{\ell+1}>p_{\ell+1}\right],$$

and its contribution to $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)$ is

$$\frac{1}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i},\,U_{\ell+1}>p_{\ell+1}+\Delta(p_{\ell+1})\right].$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

The contribution of this scenario to $f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)-f_m^\ell(\overline{p})$ is thus

$$-\frac{1}{2}\,\mathbb{P}\left[p_{\ell+1}<U_{\ell+1}\leqslant p_{\ell+1}+\Delta(p_{\ell+1}),\ \sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}\right]$$

$$=-\frac{\Delta(p_{\ell+1})}{2}\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}\right].$$

(iv) Finally, when $\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}<\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}$, its contribution to each of $f_m^\ell(\overline{p}+\Delta^{(\ell+1)}(\overline{p}))$ and $f_m^\ell(\overline{p})$ is 0, so its contribution to their difference is also 0.

Combining the contributions from the four scenarios considered above, we obtain

$$f_m^\ell\left(\overline{p}+\Delta^{(\ell+1)}(\overline{p})\right)-f_m^\ell(\overline{p})=-\frac{\Delta(p_{\ell+1})}{2}\left\{\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}+1\right]\right.$$
$$\left.+\,\mathbb{P}\left[\sum_{i=1}^{\ell}\mathbf{1}_{U_i\leqslant p_i}=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p_i}\right]\right\}. \tag{24}$$

We come back to our set-up, i.e. where $p_1=p_2=\cdots=p_m=p$ and $\Delta(p_1)=\Delta(p_2)=\cdots=\Delta(p_m)=\Delta(p)$. Setting $Z=\sum_{i=2}^{\ell}\mathbf{1}_{U_i\leqslant p}$ and $W=\sum_{i=\ell+2}^{m}\mathbf{1}_{U_i\leqslant p}$ for the sake of brevity, and recalling that $X_t(v_1)=\mathbf{1}_{U_1\leqslant p}$ and $X_t(v_{\ell+1})=\mathbf{1}_{U_{\ell+1}\leqslant p}$, we have, from (20), (21), (23), and (24),

$$\begin{aligned}
&f_m^\ell(\overline{p}+\Delta(\overline{p}))\big|_{(p+\Delta(p),p+\Delta(p),\ldots,p+\Delta(p))}-f_m^\ell(\overline{p})\big|_{(p,p,\ldots,p)}\\
&=\frac{\ell\Delta(p)}{2}\,\mathbb{P}\left[Z=X_t(v_{\ell+1})+W\right]+\frac{\ell\Delta(p)}{2}\,\mathbb{P}\left[Z=X_t(v_{\ell+1})+W-1\right]\\
&\quad-\frac{(m-\ell)\Delta(p)}{2}\,\mathbb{P}\left[X_t(v_1)+Z=W+1\right]-\frac{(m-\ell)\Delta(p)}{2}\,\mathbb{P}\left[X_t(v_1)+Z=W\right]\\
&=\frac{\Delta(p)}{2}\left\{\frac{\ell}{1-p}\,\mathbb{P}\left[Z=X_t(v_{\ell+1})+W,\,X_t(v_1)=0\right]+\ell\,\mathbb{P}\left[Z=X_t(v_{\ell+1})+W-1\right]\right.\\
&\quad\left.-(m-\ell)\,\mathbb{P}\left[X_t(v_1)+Z=W+1\right]-\frac{m-\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=W,\,X_t(v_{\ell+1})=0\right]\right\}\\
&=\frac{\Delta(p)}{2}\left\{\frac{\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=X_t(v_{\ell+1})+W,\,X_t(v_1)=0\right]+\ell\,\mathbb{P}\left[Z=X_t(v_{\ell+1})+W-1\right]\right.\\
&\quad\left.-(m-\ell)\,\mathbb{P}\left[X_t(v_1)+Z=W+1\right]-\frac{m-\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=W+X_t(v_{\ell+1}),\,X_t(v_{\ell+1})=0\right]\right\}\\
&=\frac{\Delta(p)}{2}\left\{\frac{\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=X_t(v_{\ell+1})+W\right]-\frac{\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=X_t(v_{\ell+1})+W,\,X_t(v_1)=1\right]\right.\\
&\quad+\ell\,\mathbb{P}\left[1+Z=X_t(v_{\ell+1})+W\right]-(m-\ell)\,\mathbb{P}\left[X_t(v_1)+Z=W+1\right]\\
&\quad\left.-\frac{m-\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=W+X_t(v_{\ell+1})\right]+\frac{m-\ell}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z=W+X_t(v_{\ell+1}),\,X_t(v_{\ell+1})=1\right]\right\}
\end{aligned}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$= \frac{\Delta(p)}{2}\Big\{ - \frac{(m-2\ell)}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z = X_t(v_{\ell+1})+W\right] - \frac{\ell p}{1-p}\,\mathbb{P}\left[1+Z = X_t(v_{\ell+1})+W\right]$$
$$+ \ell\,\mathbb{P}\left[1+Z = X_t(v_{\ell+1})+W\right]$$
$$- (m-\ell)\,\mathbb{P}\left[X_t(v_1)+Z = W+1\right] + \frac{(m-\ell)p}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z = W+1\right]\Big\}$$
$$= \frac{\Delta(p)}{2}\Big\{ - \frac{(m-2\ell)}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z = X_t(v_{\ell+1})+W\right] + \frac{\ell(1-2p)}{1-p}\,\mathbb{P}\left[1+Z = X_t(v_{\ell+1})+W\right]$$
$$- \frac{(m-\ell)(1-2p)}{1-p}\,\mathbb{P}\left[X_t(v_1)+Z = W+1\right]\Big\}. \tag{25}$$

In the computation above, we have made use of the mutual independence of the random variables $X_t(v_1)$, Z, $X_t(v_{\ell+1})$, and W. Note that in our set-up (i.e. where $p_1 = p_2 = \cdots = p_m = p$ and $X_t(v_i)$ are i.i.d. Bernoulli(p) for all $i \in \{1, 2, \ldots, m\}$), Z follows Binomial($\ell - 1, p$), W follows Binomial($m - \ell - 1, p$), $X_t(v_{\ell+1}) + W$ follows Binomial($m - \ell, p$), and $X_t(v_1) + Z$ follows Binomial(ℓ, p)). Moreover, Z and $X_t(v_{\ell+1}) + W$ are independent of each other, as are $X_t(v_1) + Z$ and W. We now focus on the second part of the final expression in (25), recalling from (19) that we consider $\ell \leqslant \left\lfloor \frac{m-1}{2} \right\rfloor$, which implies $\ell < m - \ell$:

$$\ell\,\mathbb{P}\left[1+Z = X_t(v_{\ell+1})+W\right] - (m-\ell)\,\mathbb{P}\left[X_t(v_1)+Z = W+1\right]$$
$$= \ell \sum_{r=0}^{\ell-1} \mathbb{P}\left[Z = r,\, X_t(v_{\ell+1})+W = r+1\right] - (m-\ell)\sum_{r=0}^{\ell-1} \mathbb{P}\left[X_t(v_1)+Z = r+1,\, W = r\right]$$
$$= \ell \sum_{r=0}^{\ell-1} \binom{\ell-1}{r} p^r (1-p)^{\ell-1-r} \binom{m-\ell}{r+1} p^{r+1}(1-p)^{m-\ell-r-1}$$
$$- (m-\ell)\sum_{r=0}^{\ell-1} \binom{\ell}{r+1} p^{r+1}(1-p)^{\ell-r-1} \binom{m-\ell-1}{r} p^r (1-p)^{m-\ell-1-r} = 0.$$

Incorporating this observation into (25), and applying (22), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}p} f_m^\ell(p) = - \frac{(m-2\ell)}{2(1-p)}\,\mathbb{P}\left[X_t(v_1)+Z = X_t(v_{\ell+1})+W\right]$$
$$= - \frac{(m-2\ell)}{2(1-p)} \sum_{r=0}^{\ell} \mathbb{P}\left[X_t(v_1)+Z = X_t(v_{\ell+1})+W = r\right]$$
$$= - \frac{m-2\ell}{2} \sum_{r=0}^{\ell} \binom{\ell}{r}\binom{m-\ell}{r} p^{2r}(1-p)^{m-1-2r},$$

which is precisely the identity in (19) that we set out to establish. From (18) and (19), we conclude that $\frac{\mathrm{d}}{\mathrm{d}p} g_m'(1/2)\big|_p > 0$ for each $p \in (0, 1)$. This completes the proof of Theorem 6. □

As explained in 4, our task now is to show that $g_m'(1/2)\big|_{p=0} < 1$ and $g_m'(1/2)\big|_{p=1} > 1$. Recall, from the very last paragraph of the proof of Theorem 5, that $f_m^\ell(1) = 0$ for each

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$\ell \leqslant \left\lfloor \frac{m-1}{2} \right\rfloor$. This, along with (18), yields

$$g'_m\left(\frac{1}{2}\right)\Big|_{p=1} = \frac{m}{2^{m-2}}\binom{m-1}{\lfloor \frac{m-1}{2} \rfloor}.$$

Comparing the values of $g'_m(1/2)|_{p=1}$ for $m = 2n+1$ and $m = 2(n+1)+1$, we obtain

$$\frac{g'_{2(n+1)+1}\left(\frac{1}{2}\right)\big|_{p=1}}{g'_{2n+1}\left(\frac{1}{2}\right)\big|_{p=1}} = \frac{2n+3}{2^{2n+1}}\binom{2n+2}{n+1}\left(\frac{2n+1}{2^{2n-1}}\binom{2n}{n}\right)^{-1} = \frac{2n+3}{2n+2} > 1,$$

for each $n \in \mathbb{N}_0$, so $\left\{g'_{2n+1}(1/2)\big|_{p=1}\right\}_{n\in\mathbb{N}}$ is a strictly increasing sequence. Likewise, comparing the values of $g'_m(1/2)\big|_{p=1}$ for $m = 2n$ and $m = 2(n+1)$, we obtain

$$\frac{g'_{2(n+1)}\left(\frac{1}{2}\right)\big|_{p=1}}{g'_{2n}\left(\frac{1}{2}\right)\big|_{p=1}} = \frac{2n+2}{2^{2n}}\binom{2n+1}{n}\left(\frac{2n}{2^{2n-2}}\binom{2n-1}{n-1}\right)^{-1} = \frac{2n+1}{2n} > 1,$$

for each $n \in \mathbb{N}$, so $\left\{g'_{2n}(1/2)\big|_{p=1}\right\}_{n\in\mathbb{N}}$ is a strictly increasing sequence. Moreover, we have $g'_3(1/2)\big|_{p=1} = 3$ and $g'_2(1/2)\big|_{p=1} = 2$. These observations, together, allow us to conclude that $g'_m(1/2)\big|_{p=1} > 1$ for every $m \in \mathbb{N}$ with $m \geqslant 2$. Finally, recall, from Remark 3, that when $p = 0$, the function $g_m(x) = 1/2$ for all $x \in [0, 1]$, so $g'_m(1/2) = 0$, for every $m \in \mathbb{N}$ with $m \geqslant 2$. As explained in 4, the proofs of 3 and 3 of Theorem 3 are now complete.

We now come to the proof of the third and last assertion made in the statement of Theorem 3. To this end, we first state and prove a lemma that will be useful in other regimes of values of (p_B, p_R) as well. The reader may refer to the green curve in Figure 1 for a pictorial representation of Lemma 2 with a specific instance of the function h.

Lemma 2. *Let h be a strictly increasing, continuous function on $[0, 1]$, with $h(0) \geqslant 0$, $h(1) \leqslant 1$, and three distinct fixed points, namely $0 \leqslant \alpha_1 < \alpha_2 < \alpha_3 \leqslant 1$. Fix any $\gamma_0 \in [0, 1]$, and let $\gamma_{n+1} = h(\gamma_n)$ for each $n \in \mathbb{N}_0$. Then the limit $\gamma = \lim_{n\to\infty} \gamma_n$ exists and equals (i) α_1 when $\gamma_0 \in [0, \alpha_2)$, (ii) α_3 when $\gamma_0 \in (\alpha_2, 1]$, and (iii) α_2 when $\gamma_0 = \alpha_2$.*

Proof. We first observe that $0 \leqslant h(x) \leqslant 1$ for each $x \in [0, 1]$. We note that the curve $y = h(x)$ lies:

(i) above the line $y = x$ for $x \in [0, \alpha_1)$ (if $\alpha_1 = 0$, which is equivalent to saying that $h(0) = 0$, this subinterval simply does not exist);

(ii) beneath the line $y = x$ for $x \in (\alpha_1, \alpha_2)$;

(iii) above the line $y = x$ for $x \in (\alpha_2, \alpha_3)$; and

(iv) beneath the line $y = x$ for $x \in (\alpha_3, 1]$ (if $\alpha_3 = 1$, which is equivalent to saying that $h(1) = 1$, this subinterval simply does not exist).

Note also that whenever the limit γ, as defined in the statement of Lemma 2, exists, it must be a fixed point of h. This is because, by continuity of the function h, we have

$$h(\gamma) = h\left(\lim_{n\to\infty} \gamma_n\right) = \lim_{n\to\infty} h(\gamma_n) = \lim_{n\to\infty} \gamma_{n+1} = \gamma. \tag{26}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

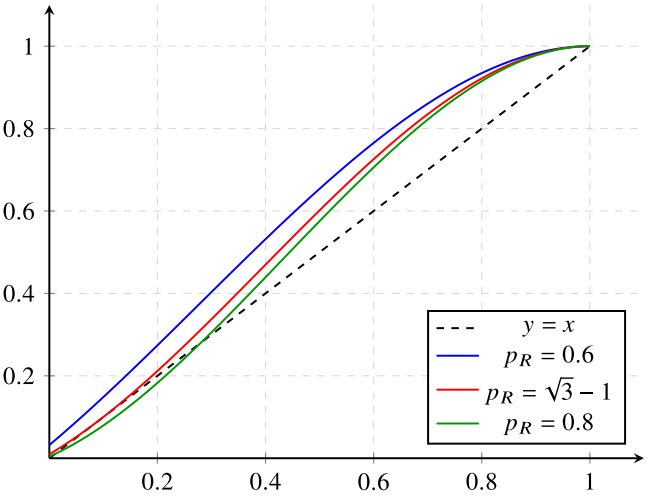


FIGURE 1. Standard plot of $y = g_3(x)$, for various values of p_R, along with $y = x$, in the $[0, 1] \times [0, 1]$ grid.

Let us consider $\gamma_0 \in [0, \alpha_1)$ (when this subinterval is non-empty, i.e. when $h(0) > 0$). The curve $y = h(x)$ lies above the line $y = x$ for $x \in [0, \alpha_1)$, so $h(\gamma_0) > \gamma_0$, and since h is strictly increasing, an iterative application of h yields $\gamma_{n+1} > \gamma_n$ for each $n \in \mathbb{N}_0$. Thus, the sequence $\{\gamma_n\}_{n\in\mathbb{N}_0}$ is strictly increasing. It is also bounded between 0 and 1, which implies that the limit $\gamma = \lim_{n\to\infty} \gamma_n$ exists. Since $\gamma_0 < \alpha_1$, the strictly increasing nature of h also ensures that $\gamma_n < \alpha_1$ for each $n \in \mathbb{N}_0$. This, in turn, implies that $\gamma \leqslant \alpha_1$ as well. Since γ must be a fixed point of h, the only possibility is that $\gamma = \alpha_1$.

When $\gamma_0 \in (\alpha_2, \alpha_3)$, we argue in the same manner as above that $\gamma = \lim_{n\to\infty} \gamma_n$ exists, and that it equals α_3. Note that in this regime, too, the curve $y = h(x)$ lies above the line $y = x$, ensuring that $\{\gamma_n\}_{n\in\mathbb{N}_0}$ is a strictly increasing sequence bounded above by α_3, and additionally is bounded below by γ_0, which is strictly greater than α_2, leading to the conclusion that $\gamma = \lim_{n\to\infty} \gamma_n$ must exist, and, by (26), must equal α_3.

Let us now consider $\gamma_0 \in (\alpha_1, \alpha_2)$. The curve $y = h(x)$ lies beneath the line $y = x$ for $x \in (\alpha_1, \alpha_2)$, so $h(\gamma_0) < \gamma_0$, and since h is strictly increasing, an iterative application of h yields $\gamma_{n+1} < \gamma_n$ for each $n \in \mathbb{N}_0$. Thus, the sequence $\{\gamma_n\}_{n\in\mathbb{N}_0}$ is strictly decreasing. It is also bounded between 0 and 1, which implies that the limit $\gamma = \lim_{n\to\infty} \gamma_n$ exists. Since $\gamma_n < \gamma_0$, and $\alpha_1 < \gamma_0$ implies, via an iterative application of the strictly increasing h, that $\alpha_1 < \gamma_n$, we conclude that $\alpha_1 < \gamma_n < \gamma_0$ for each n, so $\alpha_1 \leqslant \gamma < \gamma_0$. Since the only fixed point in $[\alpha_1, \gamma_0)$ is α_1, $\gamma = \alpha_1$ in this case.

When $\gamma_0 \in (\alpha_3, 1]$ (when this subinterval is non-empty, i.e. when $h(1) < 1$), we argue in the same manner as in the previous paragraph that $\gamma = \lim_{n\to\infty} \gamma_n$ exists, and that it equals α_3. Here, too, the curve $y = h(x)$ lies beneath the line $y = x$, ensuring that $\{\gamma_n\}_{n\in\mathbb{N}_0}$ is strictly decreasing, bounded above by γ_0, and bounded below by α_3, so $\gamma = \lim_{n\to\infty} \gamma_n$ must exist, and, by (26), must equal α_3.

Finally, if $\gamma_0 = \alpha_i$ for any $i \in \{1, 2, 3\}$ then $\gamma_n = \alpha_i$ for each $n \in \mathbb{N}$, so $\gamma = \alpha_i$ as well. This completes the proof of Lemma 2. □

As evident from Remark 2 and 3 of Theorem 3, for every choice of π_0 in (2), the sequence $\{\pi_t\}_{t\in\mathbb{N}_0}$, defined via (3), converges to the unique fixed point $1/2$ of g_m as $t \to \infty$. Therefore,

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

we need focus only on the regime $p(m) < p < 1$, where the fixed points of g_m, in the interval $[0, 1]$, are of the form α, $1/2$, and $1-\alpha$, for some $\alpha \in [0, 1/2)$, by part 3 of Theorem 3. We begin by showing that, for a fixed $m \in \mathbb{N}$ with $m \geqslant 2$ and a fixed $p \in (0, 1]$, the function g_m is strictly increasing on $[0, 1]$ (when $p = 0$, we have already seen from Remark 3 that g_m equals the constant function $1/2$). Note, from (12) and (9), that

$$\begin{aligned}
g'_m(1) &= m\,\{f_m(m) - f_m(m-1)\} \\
&= m\left\{\mathbb{P}\,[A_m > 0] + \frac{1}{2}\,\mathbb{P}\,[A_m = 0] - \mathbb{P}\,[A_{m-1} > B_1] - \frac{1}{2}\,\mathbb{P}\,[A_{m-1} = B_1]\right\} \\
&= m\Big\{1 - \frac{1}{2}\,\mathbb{P}\,[A_m = 0] - p\,\mathbb{P}\,[A_{m-1} > 1] - (1-p)\,\mathbb{P}\,[A_{m-1} > 0] - \frac{p}{2}\,\mathbb{P}\,[A_{m-1} = 1] \\
&\quad - \frac{(1-p)}{2}\,\mathbb{P}\,[A_{m-1} = 0]\Big\} \\
&= m\left\{\frac{p}{2}\,\mathbb{P}\,[A_{m-1} = 1] + \frac{1+p}{2}\,\mathbb{P}\,[A_{m-1} = 0] - \frac{1}{2}\,\mathbb{P}\,[A_m = 0]\right\} \\
&= \frac{m}{2}\left\{(m-1)p^2(1-p)^{m-2} + (1+p)(1-p)^{m-1} - (1-p)^m\right\} > 0.
\end{aligned}$$

Moreover, applying (11) to the expression in (12) and subsequently applying a change of variable to the summation (switching the index of the sum from ℓ to $m-1-\ell$), we obtain $g'_m(1-x) = g'_m(x)$, so $g'_m(0) > 0$. From Theorem 5, we know that $g''_m(x) > 0$ for each $x \in [0, 1/2)$ and $g''_m(x) < 0$ for each $x \in (1/2, 1]$, so $g'_m(x)$ is strictly increasing for $x \in [0, 1/2)$ and $g'_m(x)$ is strictly decreasing for $x \in (1/2, 1]$. Since we have shown above that both $g'_m(1)$ and $g'_m(0)$ are strictly positive, it is evident that $g'_m(x) > 0$ for all $x \in [0, 1]$, allowing us to conclude that $g_m(x)$ is strictly increasing for $x \in [0, 1]$. The rest of the last assertion made in the statement of Theorem 3 is established simply by an application of Lemma 2.

4.1. Finding the threshold probability $p(m)$ for small values of m

For small values of m, it is relatively easy to find, analytically, the value of the threshold $p(m)$ in Theorem 3. When $m = 3$, we obtain, from (9),

$$f_3(0) = \frac{1}{2}\,\mathbb{P}\,[B_3 = 0] = \frac{1}{2}(1-p)^3$$

and

$$\begin{aligned}
f_3(1) &= \mathbb{P}\,[B_2 = 0, A_1 = 1] + \frac{1}{2}\,\mathbb{P}\,[B_2 = A_1 = 1] + \frac{1}{2}\,\mathbb{P}\,[B_2 = A_1 = 0] \\
&= p(1-p)^2 + p^2(1-p) + \frac{1}{2}(1-p)^3.
\end{aligned}$$

Since (11) holds, we need not compute $f_3(2)$ and $f_3(3)$. From (5), we have

$$\begin{aligned}
g_3(x) &= f_3(0)(1-x)^3 + 3f_3(1)x(1-x)^2 + 3\,\{1 - f_3(1)\}\,x^2(1-x) + \{1 - f_3(0)\}\,x^3 \\
&= \frac{1}{2}(1-p)^3(1-2x)(1-x+x^2) + 3\left\{p(1-p)^2 + p^2(1-p) + \frac{1}{2}(1-p)^3\right\}x(1-x)(1-2x) \\
&\quad + 3x^2 - 2x^3,
\end{aligned}$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

so, upon differentiation, we obtain

$$g_3'(x) = -\frac{3}{2}(1-p)^3(1-2x+2x^2) + \frac{3}{2}(1-p)(1+p^2)(1-6x+6x^2) + 6x - 6x^2.$$

From Lemma 1, it suffices for us to find all values of p for which the derivative $g_3'(1/2) \leqslant 1$, which is equivalent to

$$\begin{aligned}
&-\frac{3}{2}(1-p)^3\left\{1-2\left(\frac{1}{2}\right)+2\left(\frac{1}{2}\right)^2\right\}+\frac{3}{2}(1-p)(1+p^2)\left\{1-6\left(\frac{1}{2}\right)+6\left(\frac{1}{2}\right)^2\right\}\\
&+6\left(\frac{1}{2}\right)-6\left(\frac{1}{2}\right)^2 \leqslant 1\\
&\iff \frac{3p^3}{2}-3p^2+3p \leqslant 1 \iff p \leqslant \frac{1}{3}(2+2^{1/3}-2^{2/3}) \approx 0.557507.
\end{aligned}$$

That the threshold $p(3)$ indeed equals 0.557507 can be deduced analytically as follows: standard techniques for finding the roots of cubic polynomials yield the only real root of the polynomial $3/2p^3 - 3p^2 + 3p - 1$ as $1/3(2 + 2^{1/3} - 2^{2/3}) \approx 0.557507$, and since this polynomial takes the value -1 at $p = 0$, it is evident that it is negative for all $p < 0.557507$ and positive for all $p > 0.557507$. Thus, when $m = 3$ and $p_B = p_R = p$, the function g_3 has a unique fixed point, which is $1/2$, in [0, 1] if and only if $p \leqslant 0.557507$. For $p > 0.557507$, the function g_3 has three distinct fixed points in [0, 1], of the form α, $1/2$, and $1 - \alpha$ for some $\alpha \in [0, 1/2)$.

When $m = 4$, we obtain, from (9),

$$f_4(0) = \frac{1}{2}\,\mathbb{P}\,[B_4 = 0] = \frac{1}{2}(1-p)^4$$

and

$$\begin{aligned}
f_4(1) &= \mathbb{P}\,[A_1 = 1, B_3 = 0] + \frac{1}{2}\,\mathbb{P}\,[A_1 = B_3 = 1] + \frac{1}{2}\,\mathbb{P}\,[A_1 = B_3 = 0]\\
&= p(1-p)^3 + \frac{3}{2}p^2(1-p)^2 + \frac{1}{2}(1-p)^4,
\end{aligned}$$

so (5), together with (11) (which also implies that $f_4(2) = 1/2$), yields

$$\begin{aligned}
g_4(x) &= f_4(0)(1-x)^4 + 4f_4(1)x(1-x)^3 + 6\cdot\frac{1}{2}x^2(1-x)^2 + 4\,\{1-f_4(1)\}\,x^3(1-x)\\
&\quad + \{1-f_4(0)\}\,x^4\\
&= \frac{1}{2}(1-p)^4(1-2x)(1-2x+2x^2) + 4\left\{p(1-p)^3 + \frac{3}{2}p^2(1-p)^2 + \frac{1}{2}(1-p)^4\right\}x(1-x)(1-2x)\\
&\quad + 3x^2 - 2x^3.
\end{aligned}$$

Differentiating this, we obtain

$$g_4'(x) = -2(1-p)^4(1-3x+3x^2) + 2(1-p)^2(1+2p^2)(1-6x+6x^2) + 6x - 6x^2,$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

so the inequality $g_4'(1/2) \leqslant 1$ becomes equivalent to

$$-2(1-p)^4\left\{1-3\left(\frac{1}{2}\right)+3\left(\frac{1}{2}\right)^2\right\}+2(1-p)^2(1+2p^2)\left\{1-6\left(\frac{1}{2}\right)+6\left(\frac{1}{2}\right)^2\right\}$$
$$+6\left(\frac{1}{2}\right)-6\left(\frac{1}{2}\right)^2\leqslant 1$$
$$\Longleftrightarrow 4p-6p^2+6p^3-\frac{5p^4}{2}\leqslant 1\Longleftrightarrow p\leqslant 0.42842.$$

That the threshold $p(4)$ indeed equals 0.42842 can be justified analytically as follows: standard techniques for solving for the roots of a quartic equation yield the only two real roots of the polynomial $4p-6p^2+6p^3-(5/2)p^4-1$ as approximately 0.42842 and 1.3300. Since the polynomial $4p-6p^2+6p^3-(5/2)p^4-1$ assumes the value -1 at $p=0$, it is evident that it is negative for all $p\in[0,\,0.42842)$ and positive for all $p\in(0.42842,\,1]$. Thus, when $m=4$ and $p_B=p_R=p$, the function g_4 has a unique fixed point, which is $1/2$, in $[0,\,1]$ if and only if $p\leqslant 0.42842$. For $p>0.42842$, the function g_4 has three distinct fixed points in $[0,\,1]$, of the form α, $1/2$, and $1-\alpha$ for some $\alpha\in[0,\,1/2)$.

5. Analysis of our learning model when $m=3$, $p_B=1$, and $p_R\in[0,\,1]$

We prove Theorem 4 in this section. We begin by computing $f_3(k)$ for $k\in\{0,\,1,\,2,\,3\}$. Since $p_B=1$, we have $A_k=k$ in (9) for each $k\in\{0,\,1,\,2,\,3\}$, so

$$f_3(0)=\frac{1}{2}\,\mathbb{P}\,[B_3=0]=\frac{1}{2}(1-p_R)^3 \text{ and } f_3(1)=\mathbb{P}\,[B_2<1]+\frac{1}{2}\,\mathbb{P}\,[B_2=1]=(1-p_R),$$

whereas $f_3(2)=f_3(3)=1$. This yields, from (5),

$$g_3(x)=\sum_{k=0}^{3}f_3(k)\binom{3}{k}x^k(1-x)^{3-k}=\frac{1}{2}(1-p_R)^3(1-x)^3+3(1-p_R)x(1-x)^2+3x^2(1-x)+x^3. \tag{27}$$

From (27), we obtain

$$g_3(x)-x=\frac{1}{2}(1-p_R)^3(1-x)^3+3(1-p_R)x(1-x)^2+3x^2(1-x)+x^3-x$$
$$=(1-x)\left\{\frac{1}{2}(1-p_R)^3(1-x)^2+3(1-p_R)x(1-x)+3x^2-x(1+x)\right\}$$
$$=(1-x)\left[\left\{\frac{1}{2}(1-p_R)^3-3(1-p_R)+2\right\}x^2+\left\{-(1-p_R)^3+3(1-p_R)-1\right\}x+\frac{1}{2}(1-p_R)^3\right]. \tag{28}$$

Let us denote by $r(x)$ the quadratic polynomial within the square brackets in (28); therefore, $r(x)$ has the roots

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$\left\{(1-p_R)^3-6(1-p_R)+4\right\}^{-1}\left[(1-p_R)^3-3(1-p_R)+1\right.$$
$$\left.\pm\sqrt{\left\{-(1-p_R)^3+3(1-p_R)-1\right\}^2-\left\{(1-p_R)^3-6(1-p_R)+4\right\}(1-p_R)^3}\right]$$
$$=\frac{(1-p_R)^3-3(1-p_R)+1\pm\sqrt{(2p_R-1)(p_R+1+\sqrt{3})\{p_R-(\sqrt{3}-1)\}}}{(1-p_R)^3-6(1-p_R)+4}.$$

First, we note that $r(0)=1/2(1-p_R)^3$ and $r(1)=1$, both of which are strictly positive when $p_R\in[0,1)$. If the leading coefficient, $1/2(1-p_R)^3-3(1-p_R)+2$, is strictly negative, which happens whenever $p_R\in[0,2-\sqrt{3})$, we see that $r(x)\to-\infty$ when $x\to\infty$ as well as when $x\to-\infty$. Therefore, the two roots of $r(x)$ will, in this case, be real and appear *outside* the interval $[0,1]$.

Let us now consider the case where the leading coefficient, $1/2(1-p_R)^3-3(1-p_R)+2$, is strictly positive, which is equivalent to focusing on $p_R\in(2-\sqrt{3},1]$. We need consider only the case where the roots of $r(x)$ are real, i.e. where the discriminant $(2p_R-1)(p_R+1+\sqrt{3})\{p_R-(\sqrt{3}-1)\}\geqslant 0$ (since g_3 has a unique fixed point in $[0,1]$, which is 1, when the roots of $r(x)$ are complex). This is true when $p_R\in(2-\sqrt{3},1/2]\cup[\sqrt{3}-1,1]$. As noted above, $r(0)$ and $r(1)$ are both strictly positive, so for the roots of $r(x)$ to exist *within* the interval $(0,1)$, we must have $\min\{r(x):x\in\mathbb{R}\}$, attained at

$$\frac{(1-p_R)^3-3(1-p_R)+1}{2\left\{\frac{1}{2}(1-p_R)^3-3(1-p_R)+2\right\}}=\frac{(1-p_R)^3-3(1-p_R)+1}{(1-p_R)^3-6(1-p_R)+4},$$

inside the interval $(0,1)$. Since the denominator has already been assumed to be positive, we now need to make sure that the numerator is positive as well, which happens whenever $p_R\in(0.6527,1]$ (note that the ratio above is already bounded above by 1, which is why we need no further conditions on p_R to ensure that $\min\{r(x):x\in\mathbb{R}\}$ is attained at a point strictly less than 1). Note that the sets $(2-\sqrt{3},1/2]\cup[\sqrt{3}-1,1]$ and $(0.6527,1]$ overlap in $[\sqrt{3}-1,1]$, which tells us that $r(x)$ has two real roots, both within the interval $(0,1)$, if and only if $p_R\in[\sqrt{3}-1,1]$.

Finally, note that when $p_R=2-\sqrt{3}$, the quadratic polynomial $r(x)$ reduces to a linear one, since the leading coefficient becomes 0, but in this case, its only root, given by $1/2(1-p_R)^3\{(1-p_R)^3-3(1-p_R)+1\}^{-1}$, will be negative since its denominator is negative. We illustrate in Figures 1 and 2 the three different scenarios for the three different regimes of values of p_R when $p_B=1$, where the blue curve is the plot of $y=g_3(x)$ corresponding to $p_R=0.6$ (which is in the regime $p_R\in[0,\sqrt{3}-1)$), the red curve is the plot of $y=g_3(x)$ corresponding to $p_R=\sqrt{3}-1\approx 0.73205$, and the green curve is the plot of $y=g_3(x)$ corresponding to $p_R=0.8$ (which is in the regime $p_R\in(\sqrt{3}-1,1)$).

When $p_R\in[0,\sqrt{3}-1)$ (see the blue curve in Figure 1), we know, by Remark 1, that no matter which $\pi_0\in[0,1]$ we choose for the distribution in (2), the limit $\pi=\lim_{t\to\infty}\pi_t$, where π_t is as defined in (3), exists and equals the unique fixed point of g_3, which is 1.

Next, we consider $p_R\in(\sqrt{3}-1,1)$ (see the green curve in Figure 1, as well as its magnified version in Figure 2, which makes it clear that the smallest fixed point of g_3, in this case, is actually strictly positive), so g_3 has three distinct fixed points in $(0,1]$ (as proved above). As in the statement of Theorem 4, let us indicate the two fixed points of g_3 that are distinct from

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

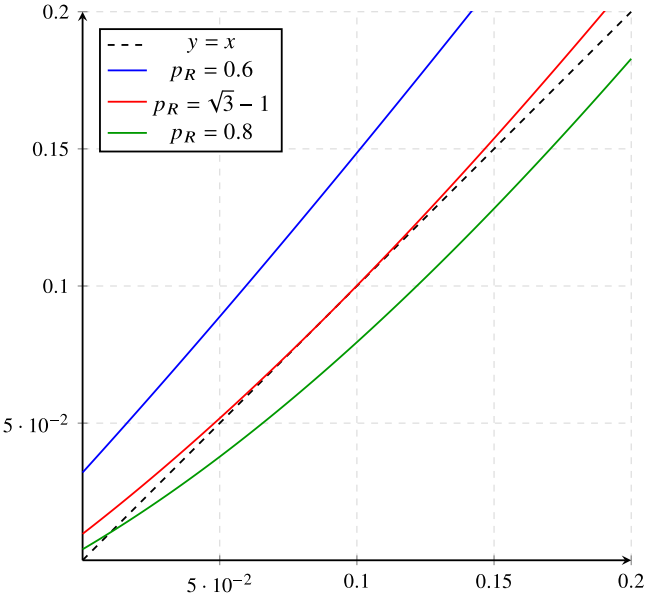


FIGURE 2. Plot of $y = g_3(x)$, for various values of p_R, along with $y = x$, magnified near the origin (0,0).

1 by α_1 and α_2, where $0 < \alpha_1 < \alpha_2 < 1$. Differentiating (27), we obtain

$$g_3'(x) = \frac{3}{2}(1 - p_R)\left(1 + 2p_R - p_R^2\right)(1 - x)^2 + 6p_R x(1 - x),$$

implying that g_3 is a strictly increasing function on [0, 1], and since $p_R < 1$, we have $g_3(0) > 0$. We thus find that g_3 is a strictly increasing, continuous function defined on [0, 1], with $g_3(0) > 0$ and $g_3(1) = 1$, and having three distinct fixed points, $0 < \alpha_1 < \alpha_2 < \alpha_3 = 1$. We are now able to conclude, letting h equal the function g_3, setting $\gamma_n = \pi_n$ for all $n \in \mathbb{N}_0$, noting the recurrence in (10) for $m = 3$, and using Lemma 2, that (i) $\pi = \alpha_1$ whenever $\pi_0 \in [0, \alpha_2)$, (ii) $\pi = 1$ whenever $\pi_0 \in (\alpha_2, 1]$, and (iii) $\pi = \alpha_2$ when $\pi_0 = \alpha_2$. Note that the case where $p_R = 1$ has already been taken care of in Theorem 3, since $p_B = p_R = 1$ in that case.

Finally, we consider $p_R = \sqrt{3} - 1$ (see the red curve in Figure 1). In this case, g_3 has precisely two distinct fixed points in [0, 1], namely $\alpha = (2/3 - 1/\sqrt{3})$ and 1, and the curve $y = g_3(x)$ merely *touches* the line $y = x$ at $x = \alpha$ instead of intersecting it. Therefore, except for the points $x = \alpha$ and $x = 1$, the curve $y = g_3(x)$ lies above the line $y = x$ throughout the entire interval [0, 1]. If $\pi_0 \in \{\alpha, 1\}$ then by (10) (with $m = 3$), obviously, $\pi_n = \pi_0$ for all n, so we have $\pi = \pi_0$ as well. Let us now choose $\pi_0 \in [0, 1] \setminus \{\alpha, 1\}$. Since the curve $y = g_3(x)$ lies above the line $y = x$ at $x = \pi_0$, we have $\pi_1 = g_3(\pi_0) > \pi_0$, and the strictly increasing nature of g_3 ensures, by an iterative application of g_3 to this inequality, that $\pi_{n+1} > \pi_n$ for each $n \in \mathbb{N}_0$, so $\{\pi_n\}_{n \in \mathbb{N}_0}$ forms a strictly increasing sequence bounded above by 1. Consequently, the limit $\pi = \lim_{n \to \infty} \pi_n$ exists. When $\pi_0 \in [0, \alpha)$, the strictly increasing nature of g_3 ensures that $0 \leqslant \pi_0 < \pi_n < \alpha$ for each $n \in \mathbb{N}$, so $0 \leqslant \pi_0 < \pi = \lim_{n \to \infty} \pi_n \leqslant \alpha$, which immediately implies that $\pi = \alpha$ (since π must be a fixed point of g_3, and the only fixed point of g_3 in $[0, \alpha]$ is α). When $\pi_0 \in (\alpha, 1)$, the strictly increasing nature of g_3 ensures $\alpha < \pi_0 < \pi_n < 1$ for each $n \in \mathbb{N}$, so $\alpha < \pi_0 < \pi = \lim_{n \to \infty} \pi_n \leqslant 1$, which immediately implies $\pi = 1$ (since π must be a fixed point of g_3, and the only fixed point of g_3 in $(\alpha, 1]$ is 1). This brings us to the end of the proof of Theorem 4.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

6. Some simulations and discussions for $m = 3$ and various values of p_B and p_R

Although a rigorous, analytical treatment of our model for *all* possible values of $(p_B, p_R) \in [0, 1]^2$ proves to be a rather intractable task for even small values of m (for any m strictly greater than 2), we include here simulations corresponding to $m = 3$ to demonstrate to the reader how the behaviour of the function g_3 (as defined in (5), with $m = 3$) differs as we vary the parameter pair (p_B, p_R) (keeping $p_B > p_R$, which causes no loss of generality since our set-up is symmetric from the perspectives of p_B and p_R). In particular, this exercise demonstrates the phenomenon of phase transition as we keep p_B fixed at a particular value in $(0, 1)$, and vary p_R in $[0, p_B)$.

From (6), letting A_k and B_{3-k} denote independent random variables where A_k follows Binomial(k, p_B) and B_{3-k} follows Binomial$(3-k, p_R)$, for each $k \in \{0, 1, 2, 3\}$, we have

$$
\begin{aligned}
f_3(0) &= \mathbb{P}[A_0 > B_3] + \frac{1}{2}\mathbb{P}[A_0 = B_3] = \frac{1}{2}\mathbb{P}[B_3 = 0] = \frac{1}{2}(1-p_R)^3,\\
f_3(1) &= \mathbb{P}[A_1 > B_2] + \frac{1}{2}\mathbb{P}[A_1 = B_2] = \mathbb{P}[A_1 = 1, B_2 = 0] + \frac{1}{2}\mathbb{P}[A_1 = B_2 = 1] + \frac{1}{2}\mathbb{P}[A_1 = B_2 = 0]\\
&= p_B(1-p_R)^2 + \frac{1}{2}p_B\binom{2}{1}p_R(1-p_R) + \frac{1}{2}(1-p_B)(1-p_R)^2\\
&= p_B(1-p_R) + \frac{1}{2}(1-p_B)(1-p_R)^2,\\
f_3(2) &= \mathbb{P}[A_2 > B_1] + \frac{1}{2}\mathbb{P}[A_2 = B_1] = \mathbb{P}[A_2 = 2] + \mathbb{P}[A_2 = 1, B_1 = 0] + \frac{1}{2}\mathbb{P}[A_2 = B_1 = 1]\\
&\quad + \frac{1}{2}\mathbb{P}[A_2 = B_1 = 0]\\
&= p_B^2 + \binom{2}{1}p_B(1-p_B)(1-p_R) + \frac{1}{2}\binom{2}{1}p_B(1-p_B)p_R + \frac{1}{2}(1-p_B)^2(1-p_R)\\
&= p_B^2 + 2p_B(1-p_B)(1-p_R) + p_Bp_R(1-p_B) + \frac{1}{2}(1-p_B)^2(1-p_R)\\
&= p_B + \frac{1}{2}(1+p_B)(1-p_B)(1-p_R),\\
f_3(3) &= \mathbb{P}[A_3 > B_0] + \frac{1}{2}\mathbb{P}[A_3 = B_0] = 1 - \mathbb{P}[A_3 = 0] + \frac{1}{2}\mathbb{P}[A_3 = 0] = 1 - \frac{1}{2}(1-p_B)^3.
\end{aligned}
$$

Combining these, we obtain

$$
\begin{aligned}
g_3(x) &= \sum_{k=0}^{3} f_3(k)\binom{3}{k}x^k(1-x)^{3-k}\\
&= \frac{1}{2}(1-p_R)^3(1-x)^3 + 3\left\{p_B(1-p_R) + \frac{1}{2}(1-p_B)(1-p_R)^2\right\}x(1-x)^2\\
&\quad + 3\left\{p_B + \frac{1}{2}(1+p_B)(1-p_B)(1-p_R)\right\}x^2(1-x) + \left\{1 - \frac{1}{2}(1-p_B)^3\right\}x^3
\end{aligned}
$$

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

$$= \frac{1}{2}(p_B + p_R)\left\{p_B^2 - 4p_Bp_R + p_R^2\right\}x^3 + \frac{3}{2}\left(p_B^2p_R - p_B^2 + 2p_Bp_R^2 - p_R^3 + p_R^2\right)x^2$$
$$+ \frac{3}{2}\left(-p_Bp_R^2 + p_B + p_R^3 - 2p_R^2 + p_R\right)x + \frac{1}{2}(1-p_R)^3, \tag{29}$$

and because of Theorem 1, we are interested in analysing the fixed points of this cubic polynomial that lie inside the interval [0, 1]. In other words, we are interested in counting the number of roots of the cubic polynomial h_3 inside [0, 1], where $h_3(x) = g_3(x) - x$ for all $x \in [0, 1]$.

We make a quick, small observation here.

Lemma 3. *For all $p_B \in [0, 1)$ and all $p_R = \theta p_B$ where $\theta \in [0, 2-\sqrt{3}) \cup (2+\sqrt{3}, \infty)$, the function g_3 has precisely one fixed point in* (0, 1).

Proof. Without loss of generality, we assume that $p_B \geqslant p_R$, so if we write $p_R = \theta p_B$, we have $\theta \in [0, 1]$. If the leading coefficient of h_3, where, from (29),

$$h_3(x) = \frac{1}{2}(p_B + p_R)\left\{p_B^2 - 4p_Bp_R + p_R^2\right\}x^3 + \frac{3}{2}\left(p_B^2p_R - p_B^2 + 2p_Bp_R^2 - p_R^3 + p_R^2\right)x^2$$
$$+ \frac{3}{2}\left(-p_Bp_R^2 + p_B + p_R^3 - 2p_R^2 + p_R - \frac{2}{3}\right)x + \frac{1}{2}(1-p_R)^3, \tag{30}$$

is strictly positive, then $\lim_{x\to\infty} h_3(x) = \infty$ and $\lim_{x\to-\infty} h_3(x) = -\infty$. Note that

$$h_3(0) = \frac{1}{2}(1-p_R)^3 > 0 \quad \text{and} \quad h_3(1) = -\frac{1}{2}(1-p_B)^3 < 0,$$

which, coupled with the previous observations, imply that the curve corresponding to the function h_3 must cross the x-axis from below at least once in the interval $(1, \infty)$, and it must cross the x-axis from above at least once in the interval $(-\infty, 0)$. Therefore, the cubic polynomial h_3 has one root in $(-\infty, 0)$ and one in $(1, \infty)$, which leaves only one root (which we already know must occur in the interval (0, 1)). Consequently, in this case, h_3 has a unique root (and hence g_3 has a unique fixed point) in (0, 1). This completes the proof. □

The analysis, however, becomes much subtler when we consider θ outside the intervals stated in Lemma 3. As such, it is well-known that should the discriminant of a cubic polynomial, with real coefficients, be strictly positive, the polynomial has three distinct real roots, whereas a strictly negative discriminant indicates the existence of two complex conjugate roots and a single real root (see, for instance, [58]). However, the coefficients in the polynomial h_3 are complicated, and functions of both p_B and p_R (or, if p_R is expressed as $p_R = \theta p_B$, of p_B and θ), and assessing the behaviour of the discriminant, which is a function of these coefficients, as p_B and θ are noth allowed to vary, is a daunting task. We have, consequently, resorted to simulations to showcase some of the patterns observed, and we hope that these will also go towards justifying a conjecture we believe holds true.

Setting $p_B = 0.4$ (so that, because of Lemma 3, we need consider only $p_R \in ((2-\sqrt{3})p_B, p_B) = (0.10718, 0.4)$), we make the following observations, using the form of the function h_3 defined in (30):

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

p_R	First root	Second root	Third root
0.11	-182.403	-2.35898	0.823719
0.12	-36.5586	-2.48484	0.819679
0.13	-18.343	-2.67089	0.815446
0.14	-10.9773	-2.97989	0.811006
0.15	-6.4548	-3.73099	0.806342
0.16	$-3.91695 + 1.84835i$	$-3.91695 - 1.84835i$	0.801438
0.17	$-3.11183 + 2.34018i$	$-3.11183 - 2.34018i$	0.796273
0.18	$-2.52476 + 2.4899i$	$-2.52476 - 2.4899i$	0.790827

Upon closer inspection, letting p_R take various values in the interval $(0.15, 0.16)$, we see that:

(i) when $p_R = 0.1525$, the roots of h_3 are approximately -4.95515, -4.54629, and 0.805139;

(ii) when $p_R = 0.1526$, the roots of h_3 are approximately $-4.73781 \pm 0.148697i$ and 0.805091;

which goes to show that a phase transition in the behaviour of h_3, in terms of the nature of its roots, happens between $p_R = 0.1525$ and $p_R = 0.1526$ when $p_B = 0.4$. It must, therefore, be the case that the discriminant corresponding to h_3 changes its sign from strictly positive to strictly negative somewhere between $p_R = 0.1525$ and $p_R = 0.1526$ when $p_B = 0.4$. Upon considering further values of p_R (such as those closer to $p_B = 0.4$), we see that for all values of $p_R < p_B = 0.4$, the function h_3 has a unique root in (0, 1) (whether by virtue of its other two real roots being strictly negative or its other two roots being complex conjugates of each other).

When we consider $p_B = 0.5$ (so that, by Lemma 3 we need consider only $p_R > (2 - \sqrt{3})0.5 \approx 0.13397$), we find that:

(i) for various values of $p_R \in (0.14, 0.207)$, the function h_3 has three distinct roots, two of which are strictly negative and the third is in (0, 1);

(ii) for various values of $p_R \in (0.208, 0.5)$, the function h_3 has two roots that are complex conjugates to each other, and one real root in (0, 1);

(iii) upon considering values of p_R in the interval $(0.207, 0.208)$, we find that for $p_R = 0.2071$, the function h_3 has roots approximately equal to -2.6654, -2.46294, and 0.885198, and for $p_R = 0.2072$, the function h_3 has roots approximately equal to $-2.5595 \pm 0.047405i$ and 0.885171, showing that a phase transition occurs between $p_R = 0.2071$ and $p_R = 0.2072$ when $p_B = 0.5$, with the discriminant corresponding to h_3 changing its sign from strictly positive to strictly negative.

Thus, the behaviour in this case is quite similar to that when we consider $p_B = 0.4$.

When we consider $p_B = 0.6$ (so that, by Lemma 3, we need consider only $p_R > (2 - \sqrt{3})0.6 \approx 0.16077$), we make the following observations:

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

p_R	First root	Second root	Third root
0.17	-49.1048	-0.831168	0.952571
0.2	-9.84522	-0.855079	0.950295
0.25	-2.8564	-1.03486	0.945944
0.3	$-0.970349 + 0.747616i$	$-0.970349 - 0.747616i$	0.940698
0.35	$-0.499685 + 0.784767i$	$-0.499685 - 0.784767i$	0.934224
0.4	$-0.217538 + 0.694856i$	$-0.217538 - 0.694856i$	0.925986
0.45	$-0.0258422 + 0.577884i$	$-0.0258422 - 0.577884i$	0.915038
0.5	$0.116512 + 0.44783i$	$0.116512 - 0.44783i$	0.899487
0.55	$0.231506 + 0.290183i$	$0.231506 - 0.290183i$	0.874565
0.59	0.265464	0.368569	0.836197

Upon closer inspection, letting p_R take values in the interval (0.25, 0.3), we see that:

(i) when $p_R = 0.2668$, the roots of h_3 are approximately -1.53883, -1.4963, and 0.944294;

(ii) when $p_R = 0.2669$, the roots of h_3 are approximately $-1.51541 \pm 0.053213i$ and 0.944284;

which shows that a phase transition happens between $p_R = 0.2668$ and $p_R = 0.2669$ when $p_B = 0.6$, with the sign of the discriminant changing from strictly positive to strictly negative. On the other hand, letting p_R take values in (0.55, 0.59), we find that:

(i) when $p_R = 0.5887$, the roots of h_3 are approximately $0.314113 \pm 0.0148406i$ and 0.838042;

(ii) when $p_R = 0.5888$, the roots of h_3 are approximately 0.313527, 0.315145, and 0.837902;

so yet another phase transition happens between $p_R = 0.5887$ and $p_R = 0.5888$ when $p_B = 0.6$, in which the sign of the discriminant changes from being strictly negative to being strictly positive. The above examples also indicate that for $p_R > p_R^{(c)}$, for some critical value $p_R^{(c)}$ that lies between 0.5887 and 0.5888, the function h_3 has multiple roots in the interval (0, 1).

From the observations made above (and from numerous other numerical examples considered that have been omitted from this paper for the sake of brevity), we are inspired to propose the following conjecture.

Conjecture 1. *There exists a critical $p_B^{(c)}$, with $p_B^{(c)} \leqslant 0.557507$, such that the function h_3 has a unique root in* (0, 1) *(equivalently, the function g_3 has a unique fixed point in* (0, 1)*) for each $p_B < p_B^{(c)}$ and each $p_R \leqslant p_B$.*

Here the value 0.557507 comes from the threshold corresponding to $m = 3$ deduced in Subsection 4.1.

Funding information

Moumanti Podder acknowledges that for the duration of the project relating to this work she has been supported by Grant No. SERB CRG/2021/006785 awarded by the Science and

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

Engineering Research Board, Government of India. Anish Sarkar acknowledges that for the duration of the project relating to this work he has been supported by CEFIPRA Project No. 6901-C, funded by Centre Franco-Indien pour la Promotion de la Recherche Avancée.

Competing interests

The authors state explicitly and without any ambiguity that they have no financial or non-financial conflict of interest to disclose that is pertinent to this work.

References

[1] Ergün, G. and Rodgers, G. J. (2002). Growing random networks with fitness. *Phys. A: Stat. Mech. Appl.* **303**(1), 261–272. https://arxiv.org/pdf/cond-mat/0103423.
[2] Ahlberg, D., Deijfen, M. and Janson, S. (2019). Competing first passage percolation on random graphs with finite variance degrees. *Random Struct. Algorithms* **55**(3), 545–559.
[3] Ahlberg, D., Griffiths, S., Janson, S. and Morris, R. (2019). Competition in growth and urns. *Random Struct. Algorithms* **54**(2), 211–227.
[4] Antunović, T., Dekel, Y., Mossel, E. and Peres, Y. (2017). Competing first passage percolation on random regular graphs. *Random Struct. Algorithms* **50**(4), 534–583.
[5] Antunović, T., Mossel, E. and Rácz, M. Z. (2016). Coexistence in preferential attachment networks. *Comb. Probab. Comput.* **25**(6), 797–822.
[6] Asavathiratham, C., Roy, S., Lesieutre, B. C. and Verghese, G. (2002). The influence model. *IEEE Control Syst. Mag.* **21**(6), 52–64.
[7] Bala, V. and Goyal, S. (1998). Learning from neighbours. *Rev. Econ. Stud.* **65**(3), 595–621.
[8] Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica* **68**(5), 1181–1229.
[9] Bandyopadhyay, A., Roy, R. and Sarkar, A. (2010). On the one dimensional "Learning from neighbours" model. *Electron. J. Probab.* **15**, 1574–1593.
[10] Banerjee, A. and Fudenberg, D. (2004). Word-of-mouth learning. *Games Econ. Behav.* **46**(1), 1–22.
[11] Baroni, E., van der Hofstad, R. and Komjáthy, J. (2015). Fixed speed competition on the configuration model with infinite variance degrees: unequal speeds. *Electron. J. Probab.* **20**(116), 1–48.
[12] Barro, R. J. (1974). Are government bonds net wealth? *J. Polit. Econ.* **82**(6), 1095–1117.
[13] Becker, G. S. and Barro, R. J. (1986). Altruism and the economic theory of fertility. *Popul. Dev. Rev.* **12**, 69–76.
[14] Berger, N., Borgs, C., Chayes, J. T. and Saberi, A. (2005). On the spread of viruses on the internet. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 301–310.
[15] Bernabeu, J. F., Calera-Rubio, J., Iñesta, J. M. and Rizo, D. (2011). Melodic identification using probabilistic tree automata. *J. New Music Res.* **40**(2), 3–103.
[16] Bezuidenhout, C. and Grimmett, G. (1990). The critical contact process dies out. *Ann. Probab.* **18**(4), 1462–1482.
[17] Bharathi, S., Kempe, D. and Salek, M. (2007). Competitive influence maximization in social networks. In *Proceedings of the Internet and Network Economics: Third International Workshop*, WINE 2007, San Diego, CA, USA, December 12–14, p. 306.
[18] Bianconi, G. and Barabási, A.-L. (2001). Competition and multiscaling in evolving networks. *Europhys. Lett.* **54**(4), 436.
[19] Bramson, M. and Griffeath, D. (1980). On the Williams-Bjerknes tumour growth model: II. *Math. Proc. Camb. Philos. Soc.* **88**, 339–357.
[20] Bramson, M. and Griffeath, D. (1981). On the Williams-Bjerknes tumour growth model: I. *Ann. Prob.* **9**, 173–185.
[21] Brown, J. J. and Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *J. Consum. Res.* **14**(3), 350–362.
[22] Carayol, A., Haddad, A. and Serre, O. (2014). Randomization in automata on infinite trees. *ACM Trans. Comput. Logic (TOCL)* **15**(3), 1–33.
[23] Castellano, C., Muñoz, M. A. and Pastor-Satorras, R. (2009). Nonlinear q-voter model. *Phys. Rev. E* **80**(4), 041129.
[24] Chatterjee, K. and Dutta, B. (2016). Credibility and strategic learning in networks. *Int. Econ. Rev.* **57**(3), 759–786.
[25] Chatterjee, K. and Xu, S. H. (2004). Technology diffusion by learning from neighbours. *Adv. Appl. Probab.* **36**(2), 355–376.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

[26] CLIFFORD, P. AND SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60**(3), 581–588.
[27] COHEN, S., KIMELFELD, B. AND SAGIV, Y. (2009). Running tree automata on probabilistic XML. In *Proceedings of the Twenty-Eighth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ACM, pp. 227–236.
[28] COMON, H., DAUCHET, M., GILLERON, R., JACQUEMARD, F., LUGIEZ, D., LÖDING, C., TISON, S. AND TOMMASI, M. (2008). Tree Automata Techniques and Applications, *Citeseer*.
[29] CONGLETON, R. D. (2004). The median voter model. In *The Encyclopedia of Public Choice*, eds. C. K. Roowley and F. Schneider, Friedrich, Springer, Dordrecht, pp. 707–712.
[30] DE MASI, A., PRESUTTI, E. AND SCACCIATELLI, E. (1989). The weakly asymmetric simple exclusion process. *Annales de l'IHP Probabilités et statistiques* **25**, 1–38.
[31] DEIJFEN, M. AND HÄGGSTRÖM, O. (2006). The initial configuration is irrelevant for the possibility of mutual unbounded growth in the two-type Richardson model. *Comb. Probab. Comput.* **15**(3), 345–353.
[32] DEIJFEN, M. AND HÄGGSTRÖM, O. (2006). Nonmonotonic coexistence regions for the two-type Richardson model on graphs. *Electron. J. Probab.* **11**, 331–344.
[33] DEIJFEN, M. AND HÄGGSTRÖM, O. (2007). The two-type Richardson model with unbounded initial configurations. *Ann. Appl. Probab.* **17**(5/6), 1639–1656.
[34] DEIJFEN, M. AND VAN DER HOFSTAD, R. (2016). The winner takes it all. *Ann. Appl. Probab.* **26**(4), 2419–2453.
[35] DOMINGOS, P. AND RICHARDSON, M. (2001). Mining the network value of customers. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 57–66.
[36] DURRETT, R. AND LIU, X.-F. (1988). The contact process on a finite set. *Ann. Probab.* **16**(3), 1158–1173.
[37] DURRETT, R. AND JUNG, P. (2007). Two phase transitions for the contact process on small worlds. *Stochastic Processes Appl.* **117**(12), 1910–1927.
[38] ELLIS, C. A. (1970). Probabilistic tree automata. In *Proceedings of the Second Annual ACM Symposium on Theory of Computing*, Association for Computing Machinery, New York, NY, USA, pp. 198–205.
[39] ELLISON, G. AND FUDENBERG, D. (1993). Rules of thumb for social learning. *J. Polit. Econ.* **101**(4), 612–643.
[40] ELLISON, G. AND FUDENBERG, D. (1995). Word-of-mouth communication and social learning. *Q. J. Econ.* **110**(1), 93–125.
[41] FUDENBERG, D. AND LEVINE, D. K. (2009). Learning and equilibrium. *Annu. Rev. Econ.* **1**(1), 385–420.
[42] GALPERTI, S. AND STRULOVICI, B. (2017). A theory of intergenerational altruism. *Econometrica* **85**(4), 1175–1218.
[43] GARET, O. AND MARCHAND, R. (2005). Coexistence in two-type first-passage percolation models. *Ann. Appl. Probab.* **15**(1A), 298–330.
[44] GARET, O. AND MARCHAND, R. (2008). First-passage competition with different speeds: positive density for both species is impossible. *Electron. J. Probab.* **13**(paper 70), 2118.
[45] GOLDENBERG, J., LIBAI, B. AND MULLER, E. (2001). Talk of the network: a complex systems look at the underlying process of word-of-mouth. *Mark. Lett.* **12**(3), 211–223.
[46] GOLDENBERG, J., LIBAI, B. AND MULLER, E. (2001). Using complex systems analysis to advance marketing theory development: modeling heterogeneity effects on new product growth through stochastic cellular automata. *Acad. Mark. Sci. Rev.* **9**(3), 1–18.
[47] GOYAL, S. (2011). Learning in networks. In *Handbook of Social Economics*, vol. 1. Elsevier, North-Holland, pp. 679–727.
[48] GOYAL, S. (2017). *Connections: An Introduction to the Economics of Networks. Princeton University Press*, Princeton, NJ.
[49] GRANOVSKY, B. L. AND MADRAS, N. (1995). The noisy voter model. *Stochastic Processes Appl.* **55**(1), 23–43.
[50] GRIFFEATH, D. (1981). The basic contact processes. *Stochastic Processes Appl.* **11**(2), 151–185.
[51] GRIMMETT, G. (1999). *What is percolation?* vol. 321, 2nd edition, Springer-Verlag, Berlin, Heidelberg.
[52] HÄGGSTRÖM, O. AND PEMANTLE, R. (1998). First passage percolation and a model for competing spatial growth. *J. Appl. Probab.* **35**(3), 683–692.
[53] HÄGGSTRÖM, O. AND PEMANTLE, R. (2000). Absence of mutual unbounded growth for almost all parameter values in the two-type Richardson model. *Stochastic Processes Appl.* **90**(2), 207–222.
[54] HASLEGRAVE, J. AND JORDAN, J. (2018). Non-convergence of proportions of types in a preferential attachment graph with three co-existing types. *Electron. Commun. Probab.* **23**(Paper No. 54), 12 p.
[55] HASLEGRAVE, J., JORDAN, J. AND YARROW, M. (2025). Competing types in preferential attachment graphs with community structure. *Electron. J. Probab.* **30**, 1–34.
[56] HOFFMAN, C. (2005). Coexistence for Richardson type competing spatial growth models. *Ann. Appl. Probab.* **15**(1B), 739–747.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

[57] HOLLEY, R. A. AND LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab*. **3**(4), 643–663.
[58] IRVING, R. S. (2004). *Integers, Polynomials, and Rings: A Course in Algebra.* Springer, New York.
[59] JORDAN, J. (2018). Preferential attachment graphs with co-existing types of different fitnesses. *J. Appl. Probab.* **55**(4), 1211–1227.
[60] KEMPE, D., KLEINBERG, J. AND TARDOS, É. (2005). Influential nodes in a diffusion model for social networks. In *Automata, Languages and Programming*, eds. L. Caires, G. F. Italiano, L. Monteiro, C. Palamidessi and M. Yung, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 3580, pp. 1127–1138.
[61] KOMAROVA, N. L. (2006). Spatial stochastic models for cancer initiation and progression. *Bull. Math. Biol.* **68**, 1573–1599.
[62] LAMBERSON, P. J. (2010). Social learning in social networks. *BE J. Theoret. Econ.* **10**(1), 00001022202193517041616.
[63] LELARGE, M. (2011). Diffusion and cascading behavior in random networks. *ACM SIGMETRICS Perform. Eval. Rev.* **39**(3), 34–36.
[64] LELARGE, M. (2012). Diffusion and cascading behavior in random networks. *Games Econom. Behav.* **75**(2), 752–775.
[65] LIGGETT, T. M. (1976). Coupling the simple exclusion process. *Ann. Probab.* **4**(3), 339–356.
[66] LIGGETT, T. M. (1997). Stochastic models of interacting systems. *Ann. Probab.* **25**(1), 1–29.
[67] LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, vol. 324, Springer-Verlag, Berlin, Heidelberg
[68] LIGGETT, T. M. (1985). *Interacting Particle Systems*, vol. 276. Springer, New York.
[69] LODEWIJKS, B. AND MAITRA, N. (2024). Long-range competition on the torus. arXiv preprint arXiv:2403.05536.
[70] LOURY, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica J. Econom. Soc.* **49**(4), 843–867.
[71] MALLICK, K. (2011). Some exact results for the exclusion process. *J. Stat. Mech.: Theory Exp.* **2011**(01), P01024.
[72] MASUDA, N., GIBERT, N. AND REDNER, S. (2010). Heterogeneous voter models. *Phys. Rev. E* **82**(1), 010103.
[73] MELE, A. (2017). A structural model of dense network formation. *Econometrica* **85**(3), 825–850.
[74] MICHEL, P., VIDAL, J.-P. AND THIBAULT, E. (2006). Intergenerational altruism and neoclassical growth models. In *Handbook of the Economics of Giving, Altruism and Reciprocity*, eds. S.-C. Kolm, and J. Mercier Ythier, Elsevier, Amsterdam, vol. 2, pp. 1055–1106.
[75] MOBILIA, M. (2003). Does a single zealot affect an infinite group of voters? *Phys. Rev. Lett.* **91**(2), 028701.
[76] MOBILIA, M. AND GEORGIEV, I. T. (2005). Voting and catalytic processes with inhomogeneities. *Phys. Rev. E* **71**(4), 046102.
[77] MOBILIA, M., PETERSEN, A. AND REDNER, S. (2007). On the role of zealotry in the voter model. *J. Stat. Mech.: Theory Exp.* **2007**(08), P08029.
[78] MOBIUS, M. AND ROSENBLAT, T. (2014). Social learning in economics. *Annu. Rev. Econ.* **6**(1), 827–847.
[79] MOSSEL, E. AND ROCH, S. (2010). Submodularity of influence in social networks: from local to global. *SIAM J. Comput.* **39**(6), 2176–2188.
[80] MUKHOPADHYAY, A., MAZUMDAR, R. R. AND ROY, R. (2020). Voter and majority dynamics with biased and stubborn agents. *J. Stat. Phys.* **181**, 1239–1265.
[81] NAMATAME, A. (2010). Diffusion and emergence in social networks. In *Intelligent Systems for Automated Learning and Adaptation: Emerging Trends and Applications*, eds. Z. T. Pan, G. A. Tsihrintzis and M. Virvou, IGI Global, Hershey, PA, pp. 231–247.
[82] NAMATAME, A., MORITA, H. AND MATSUYAMA, K. (2009). Agent-based modeling for the study of diffusion dynamics. In *SpringSim*. Citeseer.
[83] NEWMAN, M. E. J. AND WATTS, D. J. (1999). Renormalization group analysis of the small-world network model. *Phys. Lett. A* **263**(4–6), 341–346.
[84] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4**, 1–79.
[85] QUASTEL, J. (1992). Diffusion of color in the simple exclusion process. *Commun. Pure Appl. Math.* **45**(6), 623–679.
[86] RAJEWSKY, N., SANTEN, L., SCHADSCHNEIDER, A. AND SCHRECKENBERG, M. (1998). The asymmetric exclusion process: comparison of update procedures. *J. Stat. Phys.* **92**, 151–194.
[87] SANDOW, S. (1994). Partially asymmetric exclusion process with open boundaries. *Phys. Rev. E* **50**(4), 2660.
[88] SCHÜTZ, G. M. (1997). Exact solution of the master equation for the asymmetric exclusion process. *J. Stat. Phys.* **88**, 427–445.
[89] SETHI, R. AND YILDIZ, M. (2016). Communication with unknown perspectives. *Econometrica* **84**(6), 2029–2069.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press

[90] SOOD, V., ANTAL, T. AND REDNER, S. (2008). Voter models on heterogeneous networks. *Phys. Rev. E* **77**(4), 041121.
[91] SOOD, V. AND REDNER, S. (2005). Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94**(17), 178701.
[92] STACEY, A. M. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24**(4), 1711–1726.
[93] STACEY, A. M. (2001). The contact process on finite homogeneous trees. *Probab. Theory Relat. Fields* **121**(4), 551–576.
[94] TRACY, C. A. AND WIDOM, H. (2008). Integral formulas for the asymmetric simple exclusion process. *Commun. Math. Phys.* **279**(3), 815–844.
[95] TSAKAS, N. (2017). Diffusion by imitation: the importance of targeting agents. *J. Econ. Behav. Organ.* **139**, 118–151.
[96] TSAKAS, N. (2024). Optimal influence under observational learning. *Math. Soc. Sci.* **128**, 41–51.
[97] VAN DER HOFSTAD, R. (2023). Universal 'winner-takes-it-all' phenomenon in scale-free random graphs. *arXiv preprint* arXiv:2307.09911.
[98] VAN DER HOFSTAD, R. AND KOMJÁTHY, J. (2015). Fixed speed competition on the configuration model with infinite variance degrees: equal speeds. *arXiv preprint* arXiv:1503.09046.
[99] WATTS, A. (2001). A dynamic model of network formation. *Games Econ. Behav.* **34**(2), 331–341.
[100] WATTS, D. J. (2002). A simple model of global cascades on random networks. *Proc. National Acad. Sci.* **99**(9), 5766–5771.
[101] WILLIAMS, T. AND BJERKNES, R. (1972). Stochastic model for abnormal clone spread through epithelial basal layer. *Nature* **236**(5340), 19–21.
[102] YOUNG, H. P. (2009). Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. *Am. Econ. Rev.* **99**(5), 1899–1924.

https://doi.org/10.1017/apr.2026.10074 Published online by Cambridge University Press