# A learning model on the rooted regular tree

Moumanti Podder

Indian Institute of Science Education and Research (IISER) Pune
Joint work with Anish Sarkar, Indian Statistical Institute (ISI) Delhi

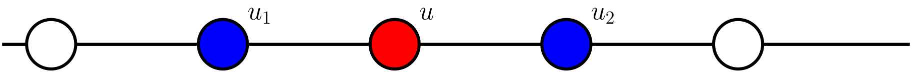
Challenges in Networks, ICTS-NETWORKS 2024
January 29, 2024

# Inspiration for this work

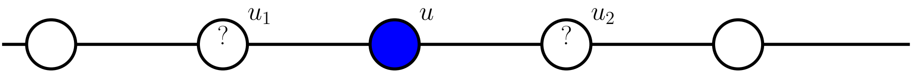
- "On the one-dimensional "learning from neighbours" model" (2010), by Antar Bandyopadhyay, Rahul Roy and Anish Sarkar.
- They studied a model of a discrete-time interacting particle system on $\mathbb{Z}$ in which infinitely many changes are allowed at each time-step.
- Imagine chameleons of two colours, blue ($B$) and red ($R$) inhabiting the cells in $\mathbb{Z}$. At each time-step, each chameleon tosses a coin (the outcomes of the coins are assumed i.i.d.).
- If the coin-toss-outcome is a success, then the chameleon retains its colour.
- Else, it looks at the colours and coin-toss-outcomes of its 2 nearest neighbours and changes its colour if and only if, out of its 2 neighbours and itself, the proportion of successes among vertices of the *other* colour exceeds the proportion of successes among vertices of its *own* colour.

# How the transitions happen

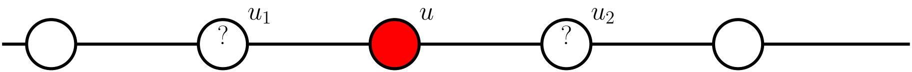
The colours of $u_1$, $u$ and $u_2$ at time-step $t$:




Let $X_t(u_1)$, $X_t(u)$ and $X_t(u_2)$ be the coin-toss-outcomes associated with $u_1$, $u$ and $u_2$ respectively at time-step $t$. When $X_t(u) = 0$, then at time-step $t+1$:



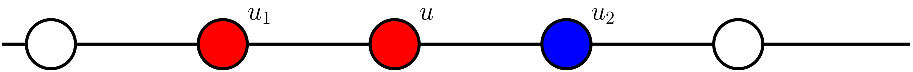

if and only if $X_t(u_1) + X_t(u_2) \geq 1$, and




if $X_t(u_1) = X_t(u_2) = 0$.

# How the transitions happen

The colours of $u_1$, $u$ and $u_2$ at time-step $t$:




Let $X_t(u_1)$, $X_t(u)$ and $X_t(u_2)$ be the coin-toss-outcomes associated with $u_1$, $u$ and $u_2$ respectively at time-step $t$. When $X_t(u) = 0$, then at time-step $t + 1$:


if and only if $X_t(u_2) = 1$, whether $X_t(u_1) = 1$ or not


if $X_t(u_2) = 0$.

# Motivation for such a model

- One motivation for studying such models arises from the notion of "social learning" studied by economists.
- "Rule of thumb for social learning" (1993), Ellison and Fudenberg introduced the concept of social learning. They studied how the speed of learning and market equilibrium is affected by social networks and other institutions governing communication among market participants.
- Two other papers that studied such models are "Learning from neighbours" (1998) and "A non-cooperative model of network formation" (2000) by Bala and Goyal.
- In "Technology diffusion by learning from neighbours" (2004), Chatterjee and Xu introduced a model consisting of particles of two types inhabiting the cells in $\mathbb{Z}$, where the type of each particle evolves with time depending on the behaviour of its neighbouring particles.

# Our model

- We let $\mathbb{T}_m$ denote the rooted tree in which each vertex has precisely $m$ children, for any $m \in \mathbb{N}$ with $m \geqslant 2$.
- We let $C_t(u) \in \{B, R\}$ denote the colour of a vertex $u \in \mathbb{T}_m$ at time-step $t$.
- We assume that $\{C_0(u) : u \in \mathbb{T}_m\}$ is a collection of i.i.d. random variables, with

$$C_0(u) = \begin{cases} B & \text{with probability } \pi_0, \\ R & \text{with probability } 1 - \pi_0 \end{cases}$$

for some $\pi_0 \in [0, 1]$.

# Questions for us all!

Interesting questions that we do not know the answers to, yet:

- What happens if $\{C_0(u) : u \in \mathbb{T}_m\}$ is not an i.i.d. collection of random variables?
- Can we at least say something if we assume that $\{C_0(u) : u \in \mathbb{T}_m\}$ is a collection of independent random variables, not necessarily identically distributed?

# Our 'general' updating / learning rule

- We define a *policy function* $f : [m] \to [0,1]$, where $[m] = \{0, 1, 2, \ldots, m\}$.
- We assume that this function is *symmetric*, i.e.

$$f(k) + f(m-k) = 1 \text{ for each } k \in [m].$$

- Let $u \in \mathbb{T}_m$, and denote its children by $u_1, u_2, \ldots, u_m$.
- Then we define

$$f(k) = \mathbf{P}\left[C_{t+1}(u) = B \,\middle|\, \sum_{i=1}^{m} \mathbf{1}_{C_t(u_i)=B} = k\right]$$

  for each $k \in [m]$.
- The update from $C_t(u)$ to $C_{t+1}(u)$ happens *independently* over *all* vertices $u$ of $\mathbb{T}_m$.

# The *absolute majority* updating / learning rule

- The update rule which we are particularly interested in is the *absolute majority* rule.
- Let $X_t(u) \in \{0,1\}$ denote the outcome of the coin-toss performed by vertex $u$ at time-step $t$. Assume $\{X_t(u) : u \in \mathbb{T}_m, t \in \mathbb{N}_0\}$ i.i.d. with each $X_t(u) \sim \text{Bernoulli}(p)$ for some $p \in [0,1]$.
- Let $\{Y_t(u) : u \in \mathbb{T}_m, t \in \mathbb{N}_0\}$ be a collection of i.i.d. Bernoulli$\left(\frac{1}{2}\right)$ random variables, independent of the former collection.
- The collection $\{X_t(u) : u \in \mathbb{T}_m\} \cup \{Y_t(u) : u \in \mathbb{T}_m\}$ is independent of $\{C_s(u) : u \in \mathbb{T}_m, s \in [t]\}$.

# The absolute majority update rule, continued

- As before, let $u_1, u_2, \ldots, u_m$ denote the children of $u$.
- Conditioned on the colours $C_t(u_1), C_t(u_2), \ldots, C_t(u_m)$ of the children at time $t$,

# The absolute majority update rule, continued

- As before, let $u_1, u_2, \ldots, u_m$ denote the children of $u$.
- Conditioned on the colours $C_t(u_1), C_t(u_2), \ldots, C_t(u_m)$ of the children at time $t$, the distribution of the colour $C_{t+1}(u)$ of the parent $u$ at time $t+1$ is defined as follows:

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

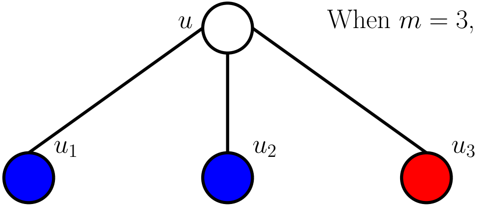
# The absolute majority update rule, continued

- As before, let $u_1, u_2, \ldots, u_m$ denote the children of $u$.
- Conditioned on the colours $C_t(u_1), C_t(u_2), \ldots, C_t(u_m)$ of the children at time $t$, the distribution of the colour $C_{t+1}(u)$ of the parent $u$ at time $t+1$ is defined as follows:

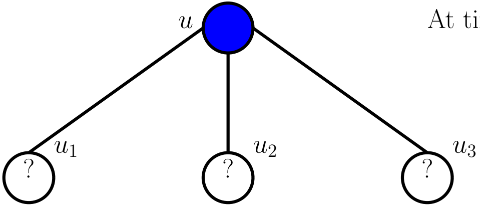
$$C_{t+1}(u) = \begin{cases} B & \text{if } \sum_{i=1}^{m} X_t(u_i)\mathbf{1}_{C_t(u_i)=B} > \sum_{i=1}^{m} X_t(u_i)\mathbf{1}_{C_t(u_i)=R}, \\ B & \text{if } \sum_{i=1}^{m} X_t(u_i)\mathbf{1}_{C_t(u_i)=B} = \sum_{i=1}^{m} X_t(u_i)\mathbf{1}_{C_t(u_i)=R} \\ & \quad \text{and } Y_t(u) = 1, \\ R & \text{otherwise.}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

- In other words, $u$ is assigned **blue** if the number of **blue** children of $u$ with successful coin-tosses exceeds the number of **red** children of $u$ with successful coin-tosses, *or* if these two numbers are exactly the same and the tie-breaking coin-toss $Y_t(u)$ results in a success.

When $m = 3$, at time-step $t$:




At time-step $t + 1$:



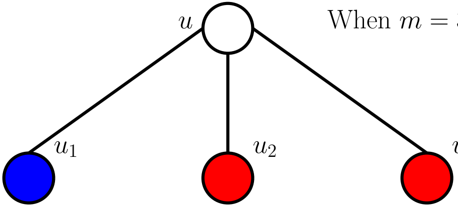

if either $X_t(u_1) + X_t(u_2) \geq 1$ and $X_t(u_3) = 0$

or $X_t(u_i) = 1$ for each $i = 1, 2, 3$

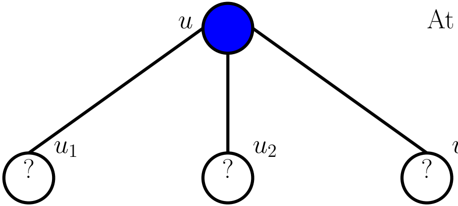
or $X_t(u_1) + X_t(u_2) = 1$ and $X_t(u_3) = 1$ and $Y_t(u) = 1$

or $X_t(u_i) = 0$ for each $i = 1, 2, 3$ and $Y_t(u) = 1$

When $m = 3$, at time-step $t$:




At time-step $t + 1$:




if either $X_t(u_1) = 1$ and $X_t(u_2) = X_t(u_3) = 0$

or $X_t(u_1) = 1$ and $X_t(u_2) + X_t(u_3) = 1$ and $Y_t(u) = 1$

or $X_t(u_i) = 0$ for each $i = 1, 2, 3$ and $Y_t(u) = 1$

# The absolute majority policy function

- We condition on $\sum_{i=1}^{m} \mathbf{1}_{C_t(u_i)=B} = k$, i.e. there being exactly $k$ blue children of $u$ at time-step $t$.
- In particular, we let $C_t(u_i) = B$ for each $i = 1, 2, \ldots, k$, and $C_t(u_i) = R$ for each $i = k+1, \ldots, m$.
- Then

$$f_{\text{abs}}(k) = \mathbf{P}\left[\sum_{i=1}^{k} X_t(u_i) > \sum_{i=k+1}^{m} X_t(u_i)\right] + \mathbf{P}\left[\sum_{i=1}^{k} X_t(u_i) = \sum_{i=k+1}^{m} X_t(u_i),\ Y_t(u) = 1\right],$$

  for each $k \in \{0, 1, \ldots, m\}$ (here, the sum over an empty set simply equals 0).
- Easy to check that the symmetry condition is satisfied.

# Recursive distributional equations

- Let the distribution of $C_t(u)$, for each $u$, be as follows:

$$C_t(u) = \begin{cases} B & \text{with probability } \pi_t, \\ R & \text{with probability } 1 - \pi_t \end{cases}$$

  for some $\pi_t \in [0, 1]$. Since we begin with an i.i.d. initial distribution, the joint distribution remains i.i.d. throughout.
- Then

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

- Thus

$$\pi_{t+1} = \mathbf{P}\left[C_{t+1}(u) = B\right] = \sum_{k=0}^{m} f_{\text{abs}}(k) \binom{m}{k} \pi_t^k (1 - \pi_t)^{m-k}.$$

## Recursive distributional equations, continued

- Let us define the function

$$g_{\mathrm{abs}}(x) = \sum_{k=0}^{m} f_{\mathrm{abs}}(k)\binom{m}{k}x^k(1-x)^{m-k},$$

  for each $k \in [m]$.
- The recurrence relation in the previous slide yields

$$\pi_{t+1} = g\left(\pi_t\right) \text{ for each } t \in \mathbb{N}_0.$$

- If $\pi := \lim_{t\to\infty} \pi_t$ exists, then it must satisfy

$$\pi = g_{\mathrm{abs}}(\pi).$$

  Thus, our task boils down to investigating all fixed points of the function $g_{\mathrm{abs}}$ in $[0,1]$.

# Fixed points of $g_{\text{abs}}$

- Since $f_{\text{abs}}$ satisfies the symmetry condition, $\pi = \frac{1}{2}$ is a fixed point for all $p \in [0,1]$ (recall that $X_t(u) \sim \text{Bernoulli}(p)$ for each $u \in \mathbb{T}_m$).
- Since $f_{\text{abs}}$ satisfies the symmetry condition, $\pi$ is a fixed point of $g_{\text{abs}}$ if and only if $1 - \pi$ is a fixed point of $g_{\text{abs}}$.

## Theorem (P., Sarkar)

*For each $m \in \mathbb{N}$ with $m \geqslant 2$, there exists $p(m) \in (0,1)$ such that the function $g_{\text{abs}}$ has a unique fixed point for all $p \leqslant p(m)$, and multiple fixed points for all $p > p(m)$.*

This result is proved by proving the following:

## Theorem (P., Sarkar)

*For each $m \in \mathbb{N}$ with $m > 4$, for each $p \in [0,1]$, the function $g_{\text{abs}}$ is strictly convex in the interval $\left[0, \frac{1}{2}\right]$ and strictly concave in the interval $\left[\frac{1}{2}, 1\right]$.*

# Idea for proving the second result

- We can write

$$g'(x) = m\mathbf{E}\left\{f_{\text{abs}}(X+1) - f_{\text{abs}}(X)\right\}, \text{ where } X \sim \text{Binomial}\,(m-1, x).$$

- Applying the same idea one more time, we obtain

$$\begin{aligned} g''(x) &= m(m-1)\mathbf{E}\left\{f_{\text{abs}}(Y+2) - 2f_{\text{abs}}(Y+1) + f_{\text{abs}}(Y)\right\}, \\ &\qquad \text{where } Y \sim \text{Binomial}\,(m-2, x) \\ &= m(m-1)\sum_{j=0}^{m-2}\left\{f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)\right\} \\ &\qquad \binom{m-2}{j}x^j(1-x)^{m-2-j}. \end{aligned}$$

- We focus on determining the sign of this last summation.

# Idea for proof, continued

- Using the symmetry condition satisfied by $f_{\text{abs}}$, we can rewrite $g''$ as

$$g''(x) = m(m-1) \sum_{j=0}^{\left\lceil \frac{m-4}{2} \right\rceil} \left\{ f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j) \right\}$$
$$\binom{m-2}{j} x^j (1-x)^j \left\{ (1-x)^{m-2-2j} - x^{m-2-2j} \right\}.$$

- If we can show that $f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j) \geqslant 0$ for each $0 \leqslant j \leqslant \left\lceil \frac{m-4}{2} \right\rceil$, then for $x \leqslant \frac{1}{2}$, we obtain $g''(x) \geqslant 0$, and for $x > \frac{1}{2}$, we have $g''(x) < 0$.

# Idea for computing $f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)$

- Recall that, if, at time $t$, we fix the colours of the children $u_1, u_2, \ldots, u_j$ to be $B$ and the colour of $u_{j+1}, u_{j+2}, \ldots, u_m$ to be $R$, then

$$f_{\text{abs}}(j) = \mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) > \sum_{i=j+1}^{m} X_t(u_i)\right] + \mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+1}^{m} X_t(u_i), Y_t(u) = 1\right].$$

- Similar expressions hold for $f_{\text{abs}}(j+1)$ and $f_{\text{abs}}(j+2)$.

# Idea for computing $f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)$

- We compute $f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)$ by considering the following different cases.
- Case 1: $\sum_{i=1}^{j} X_t(u_i) > \sum_{i=j+3}^{m} X_t(u_i) + 2$.
- Case 2: $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2$.
- Case 3: $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 1$.
- Case 4: $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i)$.
- Case 5: $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) - 1$.
- Case 6: $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) - 2$.
- Case 7: $\sum_{i=1}^{j} X_t(u_i) < \sum_{i=j+3}^{m} X_t(u_i) - 2$.

# Demonstration of analysis in one such case

- Case 2: where $\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2$. Here, the values of $X_t(u_{j+1})$ and $X_t(u_{j+2})$ act as *pivots*.
- For instance,

$$\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2 > X_t(u_{j+1}) + X_t(u_{j+2}) + \sum_{i=j+3}^{m} X_t(u_i)$$

if and only if $X_t(u_{j+1}) + X_t(u_{j+2}) < 2$, whereas

$$\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2 = X_t(u_{j+1}) + X_t(u_{j+2}) + \sum_{i=j+3}^{m} X_t(u_i)$$

if and only if $X_t(u_{j+1}) = X_t(u_{j+2}) = 1$.

# Demonstration of analysis in one such case

- Thus, the contribution of Case 2 to $f_{\text{abs}}(j)$ is

$$\mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2, (X_t(u_{j+1}), X_t(u_{j+2}))$$

$$\in \{(0,0),(0,1),(1,0)\}\Big] + \mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2,$$

$$X_t(u_{j+1}) = X_t(u_{j+2}) = 1, Y_t(u) = 1\Big].$$

- $$\sum_{i=1}^{j+1} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2 + X_t(u_{j+1}) > \sum_{i=j+3}^{m} X_t(u_i) + X_t(u_{j+2})$$

  no matter what the values of $X_t(u_{j+1})$ and $X_t(u_{j+2})$ are. Thus, the contribution of this case to $f_{\text{abs}}(j+1)$ is

$$\mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2\right].$$

# Demonstration of analysis in one such case

- The same is true for the contribution of Case 2 to $f_{\mathrm{abs}}(j+2)$.
- The contribution of Case 2 to $f_{\mathrm{abs}}(j+2) - 2f_{\mathrm{abs}}(j+1) + f_{\mathrm{abs}}(j)$ is

$$
\begin{aligned}
&\mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2\Big] - 2\,\mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2\Big] \\
&+ \mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2, (X_t(u_{j+1}), X_t(u_{j+2})) \\
&\in \{(0,0),(0,1),(1,0)\}\Big] + \mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2, \\
&X_t(u_{j+1}) = X_t(u_{j+2}) = 1, Y_t(u) = 1\Big] \\
= &-\frac{1}{2}\mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2, X_t(u_{j+1}) = X_t(u_{j+2}) = 1\Big] \\
= &-\frac{p^2}{2}\mathbf{P}\Big[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2\Big].
\end{aligned}
$$

# Putting together all the cases

- Arguing likewise, we obtain such a simplified expression for the contribution of each case to $f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)$.
- Adding them yields

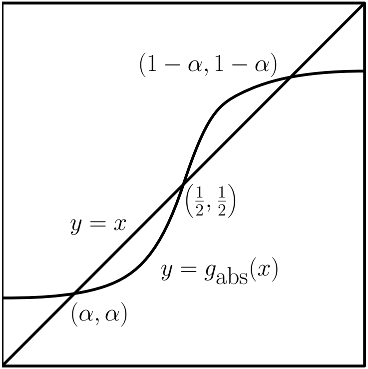
$$\frac{2}{p^2}\left\{f_{\text{abs}}(j+2) - 2f_{\text{abs}}(j+1) + f_{\text{abs}}(j)\right\} =$$

$$\mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) - 2\right] - \mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 2\right]$$

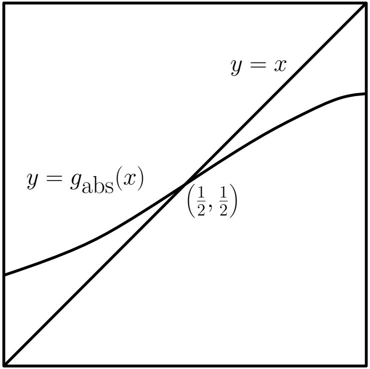
$$+\,\mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) - 1\right] - \mathbf{P}\left[\sum_{i=1}^{j} X_t(u_i) = \sum_{i=j+3}^{m} X_t(u_i) + 1\right].$$

- Recall that we only consider $0 \leqslant j \leqslant \left\lceil \frac{m-4}{2} \right\rceil$, and we show that in this range, each of the above differences is strictly positive for each $p \in [0,1]$. Incorporating into the previously shown expression for $g''(x)$, we conclude that $g''(x) > 0$ for $x \in \left[0, \frac{1}{2}\right)$ and $g''(x) < 0$ for $x \in \left(\frac{1}{2}, 1\right]$.

# To get the first result from the second




Here, $g'_{\text{abs}}(\alpha) < 1$ and $g'_{\text{abs}}\left(\frac{1}{2}\right) > 1$.




Here, $g'_{\text{abs}}\left(\frac{1}{2}\right) \leq 1$.

This allows us to conclude that $g_{\text{abs}}$ has a unique fixed point if and only if $g'\left(\frac{1}{2}\right) \leqslant 1$.

## To get the first result from the second

- It suffices to show that there exists $p(m) \in (0,1)$ such that

$$g'_{\text{abs}}\left(\frac{1}{2}\right) \leqslant 1 \text{ for each } p \in [0, p(m)],$$

  and

$$g'_{\text{abs}}\left(\frac{1}{2}\right) > 1 \text{ for each } p \in (p(m), 1].$$

- We show this by showing that $g'_{\text{abs}}\left(\frac{1}{2}\right)$ is a strictly increasing function of $p$.
- Writing

$$g'\left(\frac{1}{2}\right) = m \sum_{j=0}^{m-1} \{f_{\text{abs}}(j+1) - f_{\text{abs}}(j)\} \binom{m-1}{j} \left(\frac{1}{2}\right)^{m-1},$$

  we prove a formula for $\frac{d}{dp} f_{\text{abs}}(j)$ for each $j$ using an idea similar to proving Russo's formula.

# Further questions we could not address

- Recall that in "On the one-dimensional "learning from neighbours" model", the authors studied a policy function that takes into account *proportional majority*.
- When we try to study the same in our set-up, we run into computational complications. Is it possible to avoid that and get meaningful results?
- In "On the one-dimensional "learning from neighbours" model", the authors allowed each particle to *retain*, at time-step $t+1$, its colour from time-step $t$, unless the proportion of successes among vertices (in its neighbourhood) of the *other* colour **strictly** exceeds the proportion of successes among vertices of its *own* colour. If we incorporate this into our set-up, what happens?
- What about *general* policy functions that satisfy the symmetry condition? We could find necessary and sufficient conditions for the corresponding function $g$ to have a unique fixed point in $[0, 1]$ only for $m \in \{2, 3, 4\}$. What about higher values of $m$?

Thank you!