Infinite WARM graphs III: strong reinforcement regime

Christian Hirsch[*], Mark Holmes[†], and Victor Kleptsyn[‡]

January 13, 2021

Abstract

We study the random subgraph $\mathcal{E}_\infty$, consisting of edges reinforced infinitely often, in a reinforcement model on infinite graphs G of bounded degree. The model involves a parameter $\alpha > 0$ governing the strength of reinforcement, and Poisson firing rates λ_v at the vertices v of the graph. In [6], it was shown that for various graphs G, all connected components of $\mathcal{E}_\infty$ are finite when $\alpha \gg 1$ is sufficiently large. In [9] it was shown that infinite clusters in $\mathcal{E}_\infty$ are possible for suitably chosen G and $\alpha > 1$. In this paper, we focus on the finite connected components of $\mathcal{E}_\infty$ in the *strong reinforcement regime* ($\alpha > 1$). When α is sufficiently large, all connected components of $\mathcal{E}_\infty$ are trees.

When the firing rates λ_v are *constant*, components are trees of diameter at most 3 when α is sufficiently large. We show that there are infinitely many phase transitions as $\alpha \downarrow 1$. For instance, on the triangular lattice, increasingly large (odd) cycles appear when taking $\alpha \downarrow 1$, while on the square lattice no finite component of $\mathcal{E}_\infty$ contains a cycle for any $\alpha > 1$. Increasingly long paths and other structures appear in both lattices when taking $\alpha \downarrow 1$. In the special case where $G = \mathbb{Z}$ and $\alpha > 1$, all connected components of $\mathcal{E}_\infty$ are finite and we show that the possible cluster sizes are non-monotone in α.

An important aspect of the proofs is that on finite connected components of $\mathcal{E}_\infty$, the model behaves similarly as on finite graphs. Thus, we build on existing results concerning these processes on finite graphs, and in the course of our analysis we resolve Conjecture 1 of [8, 10] for finite graphs.

[*]University of Groningen: c.p.hirsch@rug.nl
[†]University of Melbourne: holmes.m@unimelb.edu.au
[‡]Université de Rennes 1: victor.kleptsyn@univ-rennes1.fr

1 Introduction

Pólya-type urn models are random processes where balls are repeatedly sampled from an urn, and additional balls are added depending on the colour of the sampled ball. Since their introduction in 1931 [15], generalisations of Pólya urn models have spurred a rich variety of mathematical research activity (see e.g. [14]). They are basic building blocks of competition-type probabilistic models in the fields of economics, biology and neuroscience [1, 7, 12]. A single urn is often insufficient to capture the complexity inherent in real-world applications, and consequently systems of interacting urns have gained popularity [2, 13]. In the field of neuroscience, when a neuron fires, only synapses that are connected to this neuron can be chosen to transmit the signal. Hence, Pólya models with graph-based interactions are a natural starting point for addressing one of the mechanisms of neuroplasticity: synapses that have been identified as useful in the past are more likely to be chosen in the future.

Stochastic processes described by (W, A)-reinforcement models [8] – short *WARM processes* – are a flexible framework for studying interacting Pólya urns: The strength of the reinforcement is described by a weight-function W and the interactions are determined by a sequence of subsets A_t, $t \in \mathbb{Z}_+ = \{0, 1, \dots\}$ revealing which colours are competing for selection at each step of the process. A single Pólya urn with n colours corresponds to the setting where $A_t = [n] := \{1, 2, \dots, n\}$ for every t. Included in [8, 10] is an analysis of WARM processes on finite graphs $G = (V, E)$, where the colours are the edges of the graph, and the subset A_t is the set of edges incident to an independently and randomly chosen vertex. In this setting, WARM processes describe stochastic processes of dynamically evolving integer-valued edge counts $\mathbf{N} = (N_t(e))_{e \in E, t \in \mathbb{Z}_+}$ (with e.g. $N_0(e) = 1$ for each $e \in E$).

As in [8, 10] we consider the case $W(x) = x^\alpha$ in the strong reinforcement regime and tacitly assume $\alpha > 1$ throughout the manuscript (the cases $\alpha = 1$ and $\alpha < 1$ are rather different, see e.g. [11, 5]). However, in the present paper, we focus on the natural generalisation of such models to *infinite connected graphs* (with countably many vertices). Time $t \in [0, \infty)$ is now continuous (note that $t \in \mathbb{Z}_+$ in [8, 10]). The dynamics is induced by Poisson-based firings with rates $\boldsymbol{\lambda}_V := (\lambda_v)_{v \in V}$ at the vertices V as follows:

1. Wait until the next firing of the vertex $v \in V$.

2. Choose an edge from those incident to v with probability proportional to the current count raised to the power α, i.e. choose $e \sim v$ with probability proportional to $N_{\cdot}(e)^\alpha$.

3. Increment the count of the chosen edge.

Let $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}$ denote the law of the WARM on $G = (V, E)$ with reinforcement parameter $\alpha > 1$ and firing rates $\boldsymbol{\lambda}_V$.

When the graph G is finite, the *jump process* of our model is the discrete-time WARM process studied in [8, 10], and knowledge of the finite-graph behaviour is an important ingredient in our analysis of the infinite graph setting. In the infinite setting some assumption on the firing rates and growth of the graph is required for the process to be well defined. We will assume throughout this paper that the firing rates satisfy the following condition.

Condition 1. *There exists $L > 0$ such that $0 < \lambda_v \le L$ for each $v \in V$.*

Sometimes we will restrict our attention to arguably the most interesting case where all of the firing rates are the same.

Condition 2. $\lambda_v = 1$ *for each $v \in V$.*

For convenience, we will assume that G has bounded degrees, i.e. the degrees $(\partial_x)_{x\in V}$ satisfy $\sup_x \partial_x = d$ for some $d \in \mathbb{N}$.

Condition 3. *G is a graph with bounded degrees.*

It is proved in [6, Theorem 1] that if Conditions 1 and 3 hold then the WARM process on G with rates $\boldsymbol{\lambda}_V$ is well defined. As time progresses, the number of firing events in any region grows linearly with time. However, the edge counts on specified edges need not grow at all (e.g. Rubin's construction [4] shows that for a single Pólya urn with $\alpha > 1$, only one colour is drawn infinitely often). Starting with $N_0(e) = 1$ for each $e \in E$, we investigate the random vector

$$\big(\lim_{t\to\infty} t^{-1} N_t(e)\big)_{e\in E}.$$

In particular we are interested in the random sets

$$\begin{aligned} \mathcal{E}_\infty &= \{e \in E : \sup_{t>0} N_t(e) = \infty\}, \text{ and} \\ \mathcal{E}_+ &= \{e \in E : \liminf_{t\to\infty} t^{-1} N_t(e) > 0\}. \end{aligned} \tag{1}$$

Clearly $\mathcal{E}_+ \subset \mathcal{E}_\infty$. We will prove that $\mathcal{E}_\infty = \mathcal{E}_+$ almost surely (see Proposition 1). In general the a.s. existence of *limits* (as opposed to lim inf or lim sup) in (1) is a highly non-trivial problem even on finite graphs. Success in analysing the finite setting has thus far relied on a deep connection with the fixed points of the averaged dynamics – a certain deterministic dynamical system depending on both the graph and the firing rates[1]. Let $\mathcal{S} = \mathcal{S}(G, \boldsymbol{\lambda}_V, \alpha)$ denote the set of linearly stable and critical equilibria of this deterministic dynamical system (see Section 5). The following result for finite graphs has been conjectured in [8, 10] and will be proved using a result of Tadíc [16] and some coupling arguments.

Theorem 1. *Let $G = (V, E)$ be finite, $\alpha > 1$ and $\boldsymbol{\lambda}_V \in (0, \infty)^V$. Then $(t^{-1}N_t(e))_{e\in E}$ converges almost surely to a random vector $\boldsymbol{N}$ that is supported on the set $\mathcal{S}$ of linearly stable and critical equilibria. For any linearly stable equilibrium $\boldsymbol{x} \in \mathcal{S}$, $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\boldsymbol{N} = \boldsymbol{x}) > 0$.*

In the present setting where E is infinite, although we believe that under Condition 1 the limit exists almost surely for each $\boldsymbol{\lambda}_V$, the limit of the infinite graph process does not have point masses (we do expect that the restriction of the limit to finite boxes is discrete).

Define the support $\sigma(\boldsymbol{x})$ of $\boldsymbol{x} \in \mathcal{S}$ by $\sigma(\boldsymbol{x}) = \{e \in E : x_e > 0\}$. It was proved in [10, Theorem 3] that for finite G and any $\boldsymbol{\lambda}_V \in (0, \infty)^V$, all equilibria $\boldsymbol{x}$ that are not linearly unstable (i.e. all $\boldsymbol{x} \in \mathcal{S}$) are supported on forests when $\alpha > 2$. In the case where the firing rates are constant, it was conjectured in [8] and proved in [10, Theorem 2(a)] that for any finite G and $\alpha > 25$ (not sharp), all $\boldsymbol{x} \in \mathcal{S}$ are supported on *whisker forests*, i.e. spanning graphs whose connected components are trees of diameter at most 3.

Open Problem 1. Significantly improve the bound $\alpha > 25$ from [10].

Combined with Theorem 1, this proves that on finite graphs when $\alpha > 2$ all connected components of $\mathcal{E}_+$ are trees and that they have diameter at most 3 when $\alpha > 25$. See e.g. Figure 1.

We upgrade this result to the finite components of infinite graphs.

Theorem 2. *Let G and $\boldsymbol{\lambda}_V$ satisfy Conditions 1 and 3. Then, for every edge e in a finite component of $\mathcal{E}_+$ the limit $\lim_{t\to\infty} t^{-1}N_t(e)$ exists almost surely. All finite components of $\mathcal{E}_+$ are trees if $\alpha > 2$. If also Condition 2 holds and $\alpha > 25$ then all finite components of $\mathcal{E}_+$ have diameter at most 3.*

This is not a straightforward consequence of the results on finite graphs because conditioning on $F \subset E$ being a connected component of $\mathcal{E}_+$ changes the law of the process. In particular, the law of the WARM process on G restricted to F, conditional on F being a connected component of $\mathcal{E}_+$ is not the same as the law of a WARM process on F itself.

[1] For finite graphs $\sum_{e\in E} t^{-1}N_t(e) \to \sum_{v\in V} \lambda_v$ almost surely. In such systems, by a simple time rescaling w.l.o.g. we may (and sometimes do) assume that $\sum_{v\in V} \lambda_v = 1$.

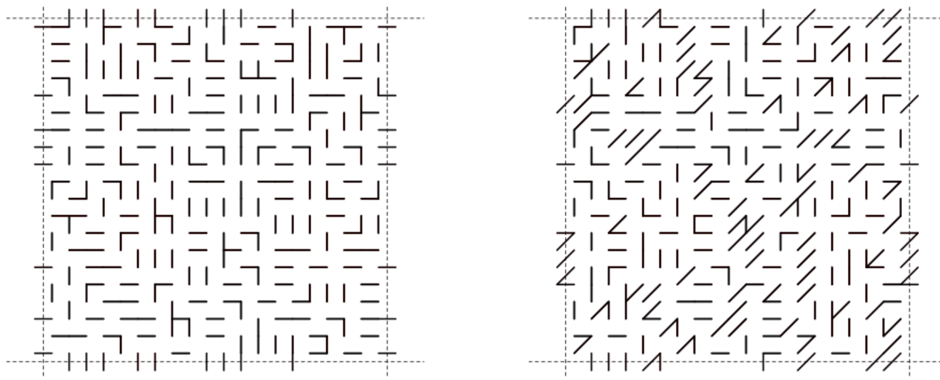


Figure 1: Simulations of $\mathcal{E}_+$ with constant firing rates on a torus (square and triangular lattices respectively) with $\alpha = 5$. All components are trees of diameter at most 3.

It was shown in [6] that for any G of bounded degree and firing rates that are bounded above, all connected components of $\mathcal{E}_+$ are a.s. finite. On the other hand for every $\alpha > 1$, [9] shows that there exists a bounded degree graph (more specifically, a regular tree) and firing rates that are bounded above (but they decrease exponentially with distance from the root) for which $\mathcal{E}_+$ contains infinite connected components.

Open Problem 2. Is it true that for any G of bounded degree, $\alpha > 1$ and $\boldsymbol{\lambda}_V$ bounded away from 0 and ∞, all components of $\mathcal{E}_+$ are a.s. finite?

Remark 1. Open Problem 2 already seems to be difficult (and worthy of study) in the settings of regular trees or $\mathbb{Z}^2$.

In the proof of Theorem 2, an important role is played by the set of edges

$$\mathcal{N} = \big\{e \in E : \sup_{t\geq 0} N_t(e) = N_0(e)\big\}$$

that are never reinforced. Roughly speaking, the parts of the WARM process on G evolving on different connected components of $\mathcal{N}^c$ evolve independently of each other (but their evolutions are governed by conditional laws).

Given a graph $G = (V, E)$ and $F \subset E$, we let $V_F = \{v \in V : (v, y) \in F \text{ for some } y \in V\} \subset V$ denote the set of vertices of F, and let $G_F = (V_F, F)$. Let $\bar{V}_F$ denote the set of vertices that are G-graph distance at most one from V_F, and $\partial F \subset E$ denote the set of edges with exactly one end-vertex in V_F.

Definition 1 (Removable edge set)**.** A set $E' \subset E$ of edges is *removable* from G if $(V, E \setminus E')$ does not have isolated vertices. That is, $V_{E\setminus E'} = V$.

For finite $F \subset E$, let $\mathcal{A}_F$ denote the event that F is a connected component of $\mathcal{E}_+$. Finally, define

$$\theta_{G,\boldsymbol{\lambda}_V,\alpha} = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{E}_+ = E)$$

which is the probability that every edge in G is used a positive proportion of the time. Note that if G is a finite star graph (i.e. every $e \in E$ is incident to a leaf $v \in V$) then $\theta_{G,\boldsymbol{\lambda}_V,\alpha} = 1$ for every α and $\boldsymbol{\lambda}_V$. In particular this holds for the graph $G = (\{v, y\}, (v, y))$ containing two vertices and a single edge. Our next main result is the following, which characterises the possible kinds of finite connected components.

Theorem 3. *Assume Conditions 1 and 3. Then, for any finite $F \subset E$,*

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{A}_F) > 0 \iff \theta_{F,\boldsymbol{\lambda}_{V_F},\alpha} > 0 \text{ and } \partial F \text{ is removable from } G.$$

Moreover, if $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{A}_F) > 0$ then the law of $(N_t(e))_{e\in F}$ under the conditional measure $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\cdot|\mathcal{A}_F)$ is absolutely continuous with respect to the law of $(N_t(e))_{e\in F}$ under the conditional measure $\mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(\cdot|\mathcal{E}_+ = F)$.

Let $\mathcal{K}_x$ denote the cluster of x in $\mathcal{E}_+$, and let

$$J_\alpha = \{k : \mathbb{P}(\mathsf{diam}(\mathcal{K}_x) = k) > 0 \text{ for some } x \in V\}$$

denote the (non-random) set of possible connected cluster sizes. Put

$$\alpha_* = \alpha(t_1) = \frac{e^{t_1}}{2 - t_1} + 1 \approx 4.4,$$

where t_1 solves $\ln\left(\frac{2}{2-e^t}\right) = \frac{te^t}{2-e^t}$. The following is our fourth main result.

Theorem 4. *Let $G = \mathbb{Z}$ and assume Condition 2. Then:*

(0) $\mathsf{diam}(0) < \infty$ *almost surely, for all $\alpha > 1$.*

(i) $1, 2 \in J_\alpha$ *for all $\alpha > 1$.*

(ii) $J_\alpha = \{1, 2, 3\}$ *for $\alpha > \alpha^*$.*

(iii) $J_\alpha = \{1, 2\}$ *for $\alpha \in (2, \alpha^*)$.*

(iv) *For every $k \geq 1$, there exists $\alpha_{2k} > 1$ such that $2k \in \cap_{\alpha<\alpha_{2k}} J_\alpha$.*

Part (0) of Theorem 4 in fact holds for graphs that are "uniformly disconnectable"[2] in the sense that there exists a constant $C > 0$ such that any finite $A \subset G$ can be disconnected from infinity by removing at most C edges from E. Such graphs include graphs that are quasi-isometric to $\mathbb{Z}$, graphs with linear volume growth, etc.

Perhaps the most striking observation from Theorem 4 is that cluster sizes are non-monotone: $\mathcal{E}_+$ admits large connected clusters for small $\alpha > 1$, clusters of size at most 2 for moderate $\alpha > 1$, and clusters up to size 3 for large $\alpha > 1$.

Open Problem 3. Is the sequence $(\alpha_{2k})_{k \in \mathbb{N}}$ (strictly) decreasing in k?

For infinite graphs with vertices of higher degree various other structures are of course possible. For example, a natural analogue of Theorem 4(i) is that star graph components are always possible (provided the boundary of the star is removable as in Theorem 3). A natural analogue of Theorem 4(iv) is that k-elongated star graphs - where each branch of the star contains k edges - are possible for $\alpha < \alpha'_k$ (an interval of length $2k$ can be thought of as an elongated star graph with 2 branches each containing k edges), see e.g. [11].

As a consequence of Theorem 3, and results about finite graphs, we obtain the following, where α_{2k} is in Theorem 4 (see also Figure 2).

Corollary 1. *Under Condition 2, with $\alpha > 1$ the following hold for the square and triangular lattices with vertex set $\mathbb{Z}^2$:*

- ■ *On the square lattice, no finite component of $\mathcal{E}_+$ contains a cycle.*
- ▲ *On the triangular lattice, for each $m \geq 1$, if $\alpha < (\cos(\pi/2n))^{-2}$ then infinitely many components of $\mathcal{E}_+$ are cycles of length $n = 2m + 1$ whose edges are used an equal limiting proportion of time. If $\alpha > (\cos(\pi/2n))^{-2}$ no components have this property.*
- ★ *For $k \geq 2$, on both lattices, infinitely many components (resp. no components) of $\mathcal{E}_+$ are simple paths of length $2k$ if $\alpha < \alpha_{2k}$ (resp. $\alpha > 2$).*

Similar results hold on other vertex transitive graphs, depending on whether they contain odd cycles. Moreover infinitely many components of $\mathcal{E}_+$ are k-elongated star graphs with central vertex having degree r if $1 < \alpha < \alpha_k^{(r)}$ provided the transitive graph contains such subgraphs. Note

[2] part (0) does not even require Condition 2

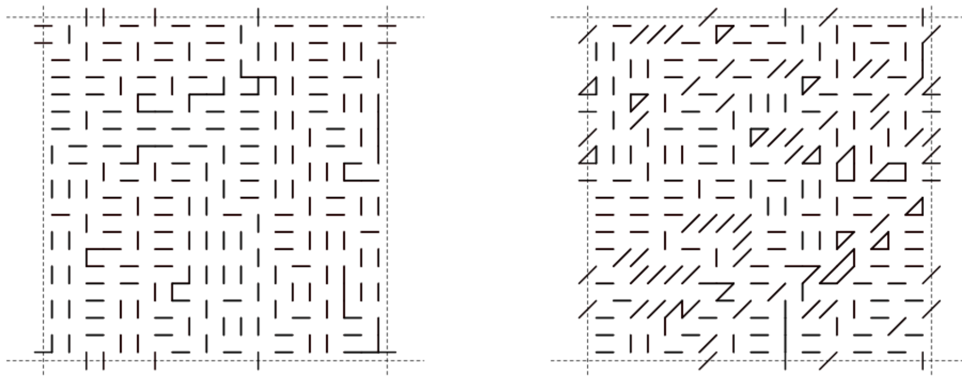


Figure 2: As in Figure 1 but with $\alpha = 1.05$. Here a large number of steps have been generated at random, and then deterministic dynamics have been used to speed up convergence (convergence is very slow when α is close to 1). In the square lattice two elongated stars of central degree 3 can be seen. In the triangular lattice cycles of length 3 and 5 are visible. Paths of length 1 and 4 are visible in both.

that Corollary 1 part ▲ shows that there are infinitely many phase transitions on the triangular lattice, as $\alpha \downarrow 1$.

Open Problem 4. Estimate the relative frequency of various types of connected components of $\mathcal{E}_+$ as a function of $\alpha > 1$ for some fixed medium or large graph. For example, estimate the relative frequency of stars of degree 3 in a length 20 torus, as a function of $\alpha > 1$.

Structure of the paper

The rest of the article is structured as follows. In Section 2, we give an explicit construction of a probability space on which our process is defined, assuming that our graph satisfies a certain condition. In Section 3, we state a number of ancillary results from which many of our main results follow. These ancillary results are proved in Section 4. Section 5 presents the basic stochastic approximation theory relevant to the evolution of a WARM on a finite graph, and proves Theorem 1. Theorems 3 and 2 are proved in Section 6. Theorem 4 and Corollary 1 are proved in Section 7, with the former proved assuming a few additional ancillary results (whose proofs are deferred to the Appendix).

2 Construction of the WARM process

In this section, we give a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which the WARM processes exist. The processes evolve on a graph $G = (V, E)$ satisfying Condition 3.

We will add some extra structure to our probability space to incorporate different processes on the same space. Let $E' \subset E$ be a possibly empty, finite removable set of edges. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we have a Poisson point process $\boldsymbol{M} = \{(X_n, T_n)\}_{n\geq 1}$ on $V \times [0, \infty)$ with intensity $\boldsymbol{\lambda}_V$, satisfying Condition 1, and a family $(U_m(x), V_m(x))_{x\in V, m\geq 1}$ of i.i.d. standard uniform random variables that are also independent of $\boldsymbol{M}$. Loosely speaking, we use $V_m(x)$ (and the current edge counts) to determine whether we choose an edge in $E^- = E \setminus E'$ or E' when the clock rings at x for the mth time. Then, we use $U_m(x)$ to determine which edge we choose from the relevant edge set. Fix an ordering $\prec$ of the edges in E^-, and also an ordering $\prec'$ of the edges in E'.

For each vector $\boldsymbol{n} = (n_e)_{e\in E} \in \mathbb{Z}_+^E$ such that $\sum_{e\in E_x} n_e > 0$ for every $x \in V$ we define a WARM process $\boldsymbol{N}^{\boldsymbol{n}} = (N_t^{\boldsymbol{n}}(e))_{e\in E, t\geq 0}$ on G as follows. The initial counts are $N_0^{\boldsymbol{n}}(e) = n_e$ for each $e \in E$. For $e_0 \in E_x \cap E^-$ define

$$R_t^{\boldsymbol{n}}(x, e_0) = \frac{N_t^{\boldsymbol{n}}(e_0)^\alpha}{\sum_{e_1\in E_x\cap E^-} N_t^{\boldsymbol{n}}(e_1)^\alpha},$$

and

$$Q_t^{\boldsymbol{n}}(x, e_0) = \sum_{e_1\in E_x\cap E^- :\, e_1\prec e_0} R_t^{\boldsymbol{n}}(x, e_1).$$

These quantities are the probabilities (conditional on the past and the event of selecting some edge from $E_x \cap E^-$) of selecting the edge e_0, and of selecting an edge less than e_0. Similarly, for $e_0 \in E_x \cap E'$ define

$$R_t'^{\boldsymbol{n}}(x, e_0) = \frac{N_t^{\boldsymbol{n}}(e_0)^\alpha}{\sum_{e_1\in E_x\cap E'} N_t^{\boldsymbol{n}}(e_1)^\alpha},$$

and

$$Q_t'^{\boldsymbol{n}}(x, e_0) = \sum_{e_1\in E_x\cap E' : e_1\prec' e_0} R_t'^{\boldsymbol{n}}(x, e_1).$$

Let $S_t^{\boldsymbol{n}}(x)$ denote the probability of selecting an edge in E^- from E_x, i.e.

$$S_t^{\boldsymbol{n}}(x) = \frac{\sum_{e_0\in E_x\cap E^-} N_t^{\boldsymbol{n}}(e_0)^\alpha}{\sum_{e_1\in E_x} N_t^{\boldsymbol{n}}(e_1)^\alpha}.$$

When the clock rings at x for the mth time at some time t, we choose a random edge $\hat{e}$ according to

$$\mathbb{1}_{\{\hat{e}=e\}} = \mathbb{1}_{\{e\in E_x\cap E^-\}}\mathbb{1}_{\left\{V_m(x)\le S^{\boldsymbol{n}}_{t-}(x)\right\}}\mathbb{1}_{\{U_m(x)\in(Q^{\boldsymbol{n}}_t(x,e),Q^{\boldsymbol{n}}_t(x,e)+R^{\boldsymbol{n}}_t(e)]\}} + \mathbb{1}_{\{e\in E_x\cap E'\}}\mathbb{1}_{\left\{V_m(x)> S^{\boldsymbol{n}}_{t-}(x)\right\}}\mathbb{1}_{\left\{U_m(x)\in(Q'^{\boldsymbol{n}}_t(x,e),Q'^{\boldsymbol{n}}_t(x,e)+R'^{\boldsymbol{n}}_t(e)]\right\}}.$$

We then increment the count of the edge $\hat{e}$, i.e. we set $N^{\boldsymbol{n}}_t(e) = N^{\boldsymbol{n}}_{t-}(e) + \mathbb{1}_{\{\hat{e}=e\}}$ for each e. One can easily check that this defines a WARM process on G with initial counts $\boldsymbol{n}$.

Remark 2. It is obvious that if G is not connected, then the WARM process on G evolves independently on different connected components of G. This is because the Poisson process is independent on disjoint components and for each $x \in V$ the set E_x only contains edges from one connected component (the component containing x).

Remark 3. Setting $n'_e = 0$ for $e \in E'$ (and $n'_e = 1$ otherwise) gives a WARM on $G' = (V, E \setminus E')$ with the Poisson point process $\boldsymbol{M}$ and the initial counts $\boldsymbol{n'} = (n'_e)_{e\in E\setminus E'}$. To see why this is the case, note that $N^{\boldsymbol{n'}}_0(e) = 0$ for each $e \in E'$ and hence $S^{\boldsymbol{n'}}_0(x) = 1$ almost surely for every x. If time T_n is the time of the mth firing at $X_n = x$ and at that time we have $N^{\boldsymbol{n'}}_{T_n^-}(e) = 0$ for each $e \in E'$, then $S^{\boldsymbol{n'}}_{T_n^-}(x) = 1$ for every $x \in V$. Hence, $V_m(x) \le S^{\boldsymbol{n'}}_{T_n^-}(x)$ almost surely, so almost surely the chosen edge is not in E'.

Remark 4. In the following, we always tacitly assume a monotone indexing of $\boldsymbol{M}$ in the sense that $T_k \le T_n$ if both $k \le n$ and $X_k = X_n$. The event $\{E' \subset \mathcal{N}\}$ equals the event

$$\bigcap_{n\ge1}\left\{V_{K(n)}(X_n) \le S^{\boldsymbol{n}}_{T_n^-}(X_n)\right\},$$

where $K(n) = |\{k \le n : X_k = X_n\}|$ denotes the number of firings at vertex X_n with index at most n. The event $\{E' \subset \mathcal{N}\}$ also equals

$$\bigcap_{v\in V: E_v\cap E'\neq\varnothing}\ \bigcap_{m\ge1}\{V_m(v) \le S^{\boldsymbol{n}}_{\tau_m(v)^-}(v)\},$$

where $\tau_m(v)$ is the time of the mth firing at v.

3 Outline of proofs and examples

Here we state a number of results that together will form the basis of the proofs of many of our main results.

The first says that if the edge weight of an edge per unit time drops too low too often, then the edge is in fact reinforced only finitely often.

Proposition 1. *Assume Conditions 1 and 3. Then, for each $e \in E$ there exists $\varepsilon(e) > 0$, depending only on α and the degrees and firing rates of its endvertices in G such that*

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}\left(\liminf_{t\to\infty} \frac{N_t(e)}{t} < \varepsilon\,,\, \sup_{t\geq 0} N_t(e) = \infty\right) = 0.$$

Proposition 2. *Assume Conditions 1 and 3. Then, for finite $E' \subset E$,*

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(E' \subset \mathcal{N}) > 0 \quad \iff \quad E' \text{ is removable from } G.$$

Moreover, if E' is finite and removable then $\mathbb{P}$-almost surely,

$$\mathbb{P}\big(E' \subset \mathcal{N} \big| \boldsymbol{M}, (U_n(x))_{n\in\mathbb{N},x\in V}\big) \geq \prod_{\substack{v\in V:\\ E_v\cap E'\neq\varnothing}} \prod_{m\geq 1} \frac{(m/|E_v|)^\alpha}{|E'\cap E_v| + (m/|E_v|)^\alpha}. \quad (2)$$

For $E' \subset E$, let $k' \in \mathbb{N} \cup \{\infty\}$ denote the number of connected components of $G \setminus E' = (V, E \setminus E')$, and let $\{G_{(i)}\}_{i\leq k'} = \{(V_{(i)}, E_{(i)})\}_{i\leq k'}$ denote the collection of connected components. If E' is finite (and G is connected) then k' is finite. If E' is finite and removable, then we let $\mu^{(i)}_{G,\boldsymbol{\lambda}_V,E'}$ denote the law of $(\boldsymbol{M}^{(i)}, \boldsymbol{N}^{(i)})$ conditional on $E' \subset \mathcal{N}$, where

$$\boldsymbol{M}^{(i)} = \boldsymbol{M} \cap (V_{(i)} \times [0,\infty))$$

is the firing process on $V_{(i)}$ and

$$\boldsymbol{N}^{(i)} = \{N_t(e)\}_{e\in E_{(i)},\, t\geq 0}$$

is the count process of edges in $E_{(i)}$.

Let $\bar{V}_{(i)}$ denote the set of vertices of graph distance at most 1 from $V_{(i)}$ and $\bar{E}_{(i)}$ denote the set of edges with both endpoints in $\bar{V}_{(i)}$. Hence, $\bar{V}_{(i)}$ includes $V_{(i)}$ and neighbours of $V_{(i)}$, while $\bar{E}_{(i)}$ includes all of $E_{(i)}$ and some edges of E'. Consider a WARM process on $\bar{G}_{(i)} = (\bar{V}_{(i)}, \bar{E}_{(i)})$ with Poisson rates $\boldsymbol{\lambda}_{V_{(i)}} = (\lambda_v)_{v\in V_{(i)}}$ for vertices in $V_{(i)}$ and 0 for vertices in $\bar{V}_{(i)} \setminus V_{(i)}$, conditional on *no edge in $\bar{E}_{(i)} \setminus E_{(i)}$ ever being chosen*, and $\mu^{o}_{\bar{G}_{(i)},\boldsymbol{\lambda}_{V_{(i)}},\bar{E}_{(i)}\cap E'}$ denote the law of that part of the process restricted to $G_{(i)}$.

Proposition 3. *Assume Conditions 1 and 3. If $E' \subset E$ is finite and removable, then conditional on $\{E' \subset \mathcal{N}\}$*

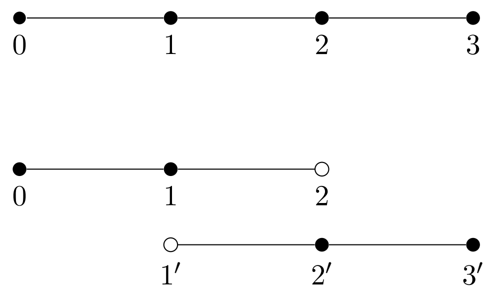


Figure 3: The path graph $G = [0, 3] \subset \mathbb{Z}$ (top) has a single removable edge $(1, 2)$. Underneath are the two graphs $\bar{G}_{(1)}$ and $\bar{G}_{(2)}$, on which we run independent WARMs (filled vertices x fire at rates λ_x and hollow vertices fire at rate 0). Conditional on the event $\{(1, 2) \in \mathcal{N}\}$, the law of the WARM process on G is the same as the joint law of two independent processes: a WARM on $\bar{G}_{(1)}$ conditional on never reinforcing edge $(1, 2)$ and a WARM on $\bar{G}_{(2)}$ conditional on never reinforcing edge $(1', 2')$.

(i) the WARM processes $\{(\boldsymbol{M}^{(i)}, \boldsymbol{N}^{(i)})\}_{i \le k'}$ *are independent, and*

(ii) the conditional distribution $\mu^{(i)}_{G,\boldsymbol{\lambda}_V,E'}$ *of the WARM process* $(\boldsymbol{M}^{(i)}, \boldsymbol{N}^{(i)})$ *given* $\{E' \subset \mathcal{N}\}$ *equals* $\mu^{o}_{\bar{G}_{(i)},\boldsymbol{\lambda}_{V_{(i)}},\bar{E}_{(i)} \cap E'}$.

The basic idea behind Proposition 3 is that if two subgraphs of G would be disconnected by removing all of the edges E', then it means that communication between the WARM sub-processes (defined by restricting the WARM process on G to each of these subgraphs) must pass through edges in E'. If those edge counts never change, then the WARM processes on the two subgraphs never communicate with (or influence) each other.

For example, when G is the path graph $[0, 3] \subset \mathbb{Z}$ of length 3, then $e = (1, 2)$ is the only removable edge. Let D_i be the event that $(1, 2)$ is never reinforced by a clock ring at i. Then, with $E' = \{(1, 2)\}$ we have $\{E' \subset \mathcal{N}\} = D_1 \cap D_2$. On the event $\{E' \subset \mathcal{N}\}$ the edge count $N_t((1, 2)) \equiv 1$ for all t, so clock rings and edge reinforcement on the right are never felt on the left and vice versa. See Figure 3.

Recall that if $F \subset E$ then ∂F is the set of edges with exactly one end-vertex in V_F. In order to prove Theorem 3 we use the fact (recall Proposition 1) that if an edge is not used a positive proportion of the time then it is used only finitely often, together with the following result.

Given a finite and removable set $E' \subset E$, let $G' = (V, E \setminus E')$. Let

$\boldsymbol{N}^- = (N_t(e))_{e \in E \setminus E', t \geq 0}$, and let $\mu^-_{G,\boldsymbol{\lambda}_V,\alpha}(\cdot) = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\boldsymbol{N}^- \in \cdot | E' \subset \mathcal{N})$ denote the conditional law of this restricted process. Let $\mu_{G',\boldsymbol{\lambda}_V,\alpha}$ denote the law of the WARM process on G'.

Proposition 4. *Let $E' \subset E$ be finite and removable. Then, $\mu^-_{G,\boldsymbol{\lambda}_V,\alpha}$ and $\mu_{G',\boldsymbol{\lambda}_V,\alpha}$ are mutually absolutely continuous.*

Proposition 5. *Assume Conditions 1 and 3. Let $E' \subset E$ be removable and finite, and $R \in \sigma(N_t(e) : t \geq 0, e \notin E')$. Then,*

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, E' \subset \mathcal{E}^c_\infty) > 0 \iff \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, E' \subset \mathcal{N}) > 0.$$

Proposition 6. *Let G be finite and $E' \subset E$ be removable. Then, $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{E}_+ = E \setminus E') > 0$ if and only if $\theta_{G_i,\boldsymbol{\lambda}_{V_i}}(\alpha) > 0$ for each $i \leq k'$.*

For $v \in V$, let $\mathcal{E}_v$ denote the connected component of v in $\mathcal{E}_+$. The WARM model on $G = (V, E)$ under Condition 2 can be defined on a probability space (similar to the construction in Section 2) in terms of a family $\big((T_{x,i}, U_{x,i})_{i \in \mathbb{N}}\big)_{x \in V}$, of random variables that are i.i.d. over sites $x \in V$: the $T_{x,i}$ are Poisson firing times at $x \in V$, and we use $U_{x,i}$ to pick the edge that is reinforced upon the ith firing at x. This "environment" is shift/translation invariant on a transitive graph, so an application of the ergodic theorem immediately yields the following result, which says that for a transitive graph G, almost surely, for any given finite set of edges F there are either no or infinitely many connected components that are translations of F.

Proposition 7. *Let $G = (V, E)$ be a transitive graph with distinguished vertex $o \in V$, and assume Condition 2. If $F \ni o$ is a finite connected subset of E and $\mathbb{P}_{G,1,\alpha}(\mathcal{A}_F) > 0$ then*

$$\mathbb{P}_{G,1,\alpha}(\mathcal{E}_v \textit{ is isomorphic to } F \textit{ for infinitely many } v) = 1.$$

4 Proof of Propositions 1-6

To prove Proposition 1, we begin by introducing a bivariate Markov chain coupled with our WARM process. Loosely speaking, it captures the joint evolution of the maximum weight of an edge incident to a vertex together with the weight of a specific edge.

Let $(W_t)_{t \geq 0} = ((W^{[1]}_t, W^{[2]}_t))_{t \geq 0}$ denote a discrete-time Markov chain with state space $\mathbb{N} \times \mathbb{Z}_+$, and transition probabilities as follows, where $r \in \mathbb{N}$, and $\theta \in (0,1)$

$$(n,s) \mapsto \begin{cases} (n, s+1) & \text{with probability } \left(\frac{r}{n}\right)^{\alpha}, \\ (n+1, s) & \text{with probability } \theta\left(1 - \left(\frac{r}{n}\right)^{\alpha}\right), \\ (n,s) & \text{otherwise.} \end{cases}$$

Let the initial state be $(w^{[1]}, w^{[2]})$, and let $\zeta = r/w^{[1]}$.

Lemma 1. *For each $\alpha > 1$ there exists $c_\alpha > 0$ such that if $\zeta < 1/2$ then*

$$\mathbb{P}\big(\sup_{t\geq 1} W_t^{[2]} - w^{[2]} \geq r/3\big) < \frac{c_\alpha}{\theta}\zeta^{\alpha-1}.$$

Proof. For $m \geq 1$, let $k_m = \inf\{t : W_t^{[1]} = m\}$ denote the first hitting time of m, so that $W_t^{[1]} = m$ if and only if $k_m \leq t \leq k_{m+1} - 1$. Let $K(m) = k_{m+1} - k_m$ denote the number of steps that $W_t^{[1]}$ stays in state m. Then, $K(m) \sim \mathsf{Geometric}(\gamma(m))$, where

$$\gamma(m) := \theta\left(1 - \left(\frac{r}{m}\right)^{\alpha}\right).$$

We can now calculate the expectation of $\Delta := \sup_{n\geq 1} W_n^{[2]} - W_0^{[2]}$. Indeed,

$$\Delta = \sum_{m \geq w^{[1]}} (W_{k_{m+1}}^{[2]} - W_{k_m}^{[2]}), \tag{3}$$

hence equality also holds for the expectations.

Now, conditionally on $K(m)$ the difference $W_{k_{m+1}}^{[2]} - W_{k_m}^{[2]}$ is a sum of $K(m) - 1$ independent Bernoulli random variables with parameter

$$\rho(m) := \frac{\left(\frac{r}{m}\right)^{\alpha}}{1 - \gamma(m)}.$$

Thus,

$$\begin{aligned} \mathbb{E}\big[W_{k_{m+1}}^{[2]} - W_{k_m}^{[2]}\big] &= \mathbb{E}\big[\mathbb{E}[W_{k_{m+1}}^{[2]} - W_{k_m}^{[2]} \mid K(m)]\big] \\ &= \rho(m)\mathbb{E}[K(m) - 1] = \rho(m)\left[\frac{1}{\gamma(m)} - 1\right] \\ &= \frac{\left(\frac{r}{m}\right)^{\alpha}}{\gamma(m)} = \frac{\left(\frac{r}{m}\right)^{\alpha}}{\theta\left(1 - \left(\frac{r}{m}\right)^{\alpha}\right)}. \end{aligned}$$

Now, since in the sum (3), $m \geq w^{[1]}$, we see that $\frac{r}{m} \leq \zeta$ for such m. If $\zeta \leq 1/2$, then

$$\mathbb{E}[\Delta] \leq \sum_{m \geq w^{[1]}} \frac{\left(\frac{r}{m}\right)^\alpha}{\theta\left(1 - \left(\frac{r}{m}\right)^\alpha\right)} \leq \frac{2}{\theta} \sum_{m \geq w^{[1]}} \left(\frac{r}{m}\right)^\alpha \leq \frac{c'_\alpha r^\alpha}{\theta (w^{[1]})^{\alpha-1}},$$

where c'_α only depends on α.

Finally, by Markov's inequality,

$$\mathbb{P}(\Delta > r/3) \leq \frac{3}{r}\mathbb{E}[\Delta] \leq \frac{3c'_\alpha}{\theta}\left(\frac{r}{w^{[1]}}\right)^{\alpha-1} = \frac{3c'_\alpha}{\theta}\zeta^{\alpha-1},$$

as claimed. ■

We are now ready to prove Proposition 1 via Lemma 1.

Proof of Proposition 1. Fix $e_0 = \{x, y\} \in E$. Then, by symmetry, it suffices to show that (for sufficiently small ε) almost surely on the event

$$O_{e_0}(\varepsilon) := \left\{ \liminf_{t\to\infty} t^{-1} N_t(e_0) < \varepsilon \right\}$$

the edge e_0 is reinforced only finitely often from the vertex x. Almost surely there exists $\tau > 0$ such that for all $t > \tau$ the number of Poisson firings at x and y at time t is between $t/2$ and $2t$ times their respective firing rates. Then, at each time $t > \tau$,

$$M_t(x) := \max_{e \in E_x} N_t(e) > \frac{\lambda_x t}{2d_x},$$

and similarly for y. It follows that almost surely on the event $O_{e_0}(\varepsilon)$ there are infinitely many times $t_0, t_1, \ldots$ at which $t_i^{-1} N_{t_i}(e_0) < \varepsilon$ and $M_{t_i}(x) > \lambda_x t_i/(2d_x)$ and $M_{t_i}(y) > \lambda_y t_i/(2d_y)$. Taking $\varepsilon < \lambda_x/(4d_x)$ ensures that $2N_{t_i}(e_0) \leq M_{t_i}(x)$ for each i.

Let $t_i > \tau$ be such a time. Then, at any time $s > t_i$ such that $N_s(e_0) < 2N_{t_i}(e_0)$ and a firing occurs at x, there is probability at most $(2N_{t_i}(e_0)/M_s(x))^\alpha$ that the WARM process chooses to reinforce edge e_0 and probability at least $(1 - (2N_{t_i}(e_0)/M_s(x))^\alpha)/d_x$ of choosing to reinforce a specific edge $e \in E_x$ with maximal weight $M_s(x)$ (so $e \neq e_0$). Let $r = 2N_{t_i}(e_0)$ and $w^{[1]} = M_{t_i}(x)$, $w^{[2]} = 0$, and $\theta = \frac{1}{d_x}$. Let $\zeta^* < 1/2$ be sufficiently small (depending only on α, θ) so that $c_\alpha(\zeta^*)^{\alpha-1}/\theta < 1/3$, and choose $\varepsilon < \frac{\zeta^*}{4}\left(\frac{\lambda_x}{d_x} \wedge \frac{\lambda_y}{d_y}\right)$ so that

$$\frac{r}{w^{[1]}} = \frac{2N_{t_i}(e_0)}{M_{t_i}(x)} \leq 4\varepsilon\frac{d_x}{\lambda_x} < \zeta^*.$$

Then, according to Lemma 1, with this choice of ε there is probability at least 2/3 the WARM process observed on the neighbours of x reinforces e_0 from x no more than $N_{t_i}(e_0)/3$ times after time t_i. We therefore see this occur after at most $I \sim \mathsf{Geometric}(2/3)$ number of attempts. Thus, $(N_t(e_0))_{t\geq 0}$ is bounded almost surely on the event $O_{e_0}(\varepsilon)$. ■

Proof of Proposition 2. First, suppose that $E' \subset E$ is not removable. By definition, $G \setminus E'$ contains an isolated vertex $x \in V$. Since $\lambda_x > 0$ we have $\#(\boldsymbol{M} \cap (\{x\} \times [0,\infty)) = \infty$ almost surely, at least one of $\{N_t(e)\}_{e\in E_x}$ has to diverge to ∞. Thus, $\mathbb{P}(E' \subset \mathcal{N}) = 0$.

Next, once (2) is verified, it is now an easy exercise to show that (if E' is removable and finite) the right hand side of (2) is strictly positive since $\{v \in V : E_v \cap E' \neq \varnothing\}$ is finite and $\alpha > 1$. Hence, if E' is finite and removable then $\mathbb{P}(E' \subset \mathcal{N}) > 0$. Hence, it remains to prove (2).

Let $\boldsymbol{U} = (U_n(x))_{n\in\mathbb{N},x\in V}$. Then, the left-hand side of (2) is equal to the limit as $r \to \infty$ of

$$\sum_{v\in V}\sum_{k\in\mathbb{N}} \mathbb{1}_{\{X_r=v\}}\mathbb{1}_{\{K(r)=k\}}\mathbb{P}\big(\cap_{j\leq r}\{V_{K(j)}(X_j) \leq S^{\boldsymbol{n}}_{T_j^-}(X_j)\}\big|\boldsymbol{M},\boldsymbol{U}\big). \tag{4}$$

Write the probability as

$$\begin{aligned}
&\mathbb{P}\big(\cap_{j\leq r}\{V_{K(j)}(X_j) \leq S^{\boldsymbol{n}}_{T_j^-}(X_j)\}\big|\boldsymbol{M},\boldsymbol{U}\big)\\
&= \mathbb{P}\big(V_{K(r)}(X_r) \leq S^{\boldsymbol{n}}_{T_r^-}(X_r)\big|\boldsymbol{M},\boldsymbol{U},\cap_{j\leq r-1}\{V_{K(j)}(X_j) \leq S^{\boldsymbol{n}}_{T_j^-}(X_j)\}\big) \qquad (5)\\
&\quad \times \mathbb{P}\big(\cap_{j\leq r-1}\{V_{K(j)}(X_j) \leq S^{\boldsymbol{n}}_{T_j^-}(X_j)\}\big|\boldsymbol{M},\boldsymbol{U}\big).
\end{aligned}$$

On the event $\{X_r = v\} \cap \{K(r) = k\}$, the current firing is the kth firing of the vertex v. If the vertex v is not incident to an edge in E', then $S^{\boldsymbol{n}}_{T_r^-}(X_r) = 1$, and therefore the probability in (5) is 1. Otherwise v is incident to an edge in E', and at least one edge in E_v has count at least $1 + (k-1)/|E_v| \leq k/|E_v|$. On the event that previous firings have not reinforced an edge in E', at least one edge in $E_v \setminus E'$ has count at least $k/|E_v|$, and moreover that $\sum_{e\in E_v\cap E'} N_{T_r^-}(e) = |E' \cap E_v|$ and hence

$$S^{\boldsymbol{n}}_{T_r^-}(X_r) \geq \frac{(k/|E_v|)^\alpha}{|E' \cap E_v| + (k/|E_v|)^\alpha}.$$

Thus, if $E_v \cap E' \neq \varnothing$, then (5) is at least

$$\mathbb{P}\left(V_{K(r)}(X_r) \leq \frac{(k/|E_v|)^\alpha}{|E' \cap E_v| + (k/|E_v|)^\alpha}\big|\boldsymbol{M},\boldsymbol{U},\cap_{j\leq r-1}\{V_{K(j)}(X_j) \leq S^{\boldsymbol{n}}_{T_j^-}(X_j)\}\right).$$

But $V_{K(r)}(X_r)$ is independent of the conditioning, so this probability is equal to $\frac{(k/|E_v|)^\alpha}{|E' \cap E_v| + (k/|E_v|)^\alpha}$. Thus, (4) is bounded below by

$$\sum_{v \in V} \sum_{k \in \mathbb{N}} \mathbb{1}_{\{X_r = v\}} \mathbb{1}_{\{K(r)=k\}} \left[\mathbb{1}_{\{E_v \cap E' = \varnothing\}} + \mathbb{1}_{\{E_v \cap E' \neq \varnothing\}} \frac{(k/|E_v|)^\alpha}{|E' \cap E_v| + (k/|E_v|)^\alpha} \right]$$
$$\times \mathbb{P}\left(\cap_{j \le r-1} \{ V_{K(j)}(X_j) \le S^{\boldsymbol{n}}_{T_j^-}(X_j) \} \big| \boldsymbol{M}, \boldsymbol{U} \right).$$

Proceeding inductively we get that (4) is bounded below by

$$\sum_{\substack{v_1,\dots,v_r \\ \in V}} \sum_{\substack{k_1,\dots,k_r \\ \in \mathbb{N}}} \prod_{j \le r} \mathbb{1}_{\{X_j = v_j\}} \mathbb{1}_{\{K(j)=k_j\}} \Big[\mathbb{1}_{\left\{E_{v_j} \cap E' = \varnothing\right\}}$$
$$+ \mathbb{1}_{\left\{E_{v_j} \cap E' \neq \varnothing\right\}} \frac{(k_j/|E_{v_j}|)^\alpha}{|E' \cap E_v| + (k_j/|E_{v_j}|)^\alpha} \Big].$$

Letting $K(n,v) = |\{k \le n : X_k = v\}|$ denote the number of firings at vertex v with index at most n, we see that this equals

$$\prod_{v : E_v \cap E' \neq \varnothing} \prod_{j \le K(r,v)} \frac{(j/|E_v|)^\alpha}{|E'| + (j/|E_v|)^\alpha}.$$

Taking the limit as $r \to \infty$ and using the fact that $K(r,v) \to \infty$ a.s. for every v verifies (2). ■

Proof of Proposition 3. Let $E' \subset E$ be finite and removable. Recall the probability space of Section 2. Then, $\mathbb{P}(E' \subset \mathcal{N}) > 0$ by Proposition 2. Let $\bar{\boldsymbol{N}}^{(i)} = (\bar{N}_t^{(i)}(e))_{e \in \bar{E}_{(i)}, t \ge 0}$ denote the WARM process on $\bar{G}_{(i)}$ induced by the construction on our probability space, but with $\bar{\boldsymbol{\lambda}}_V = (\bar{\lambda}_x)_{x \in V}$ defined by $\bar{\lambda}_x = \lambda_x \mathbb{1}_{\left\{x \notin \bar{V}_{(i)} \setminus V_{(i)}\right\}}$. In other words, we use the same independent uniform random variables (defined site by site) and edge ordering, and we use the Poisson point process $\mathbf{M}^{(i)}$, except that we remove all firings at sites in $\bar{V}_{(i)} \setminus V_{(i)}$. Here, edges in $\bar{E}_{(i)} \setminus E_{(i)}$ can only be reinforced from a firing at the endvertex in $V_{(i)}$ (since there are no firings at the other endvertex).

Introduce the event

$$K_i^* := \left\{ V_m(x) \le w_x(\{\bar{N}^{(i)}_{T_m(x)-}(e)\}_{e \in E_x \setminus E'}) \text{ for all } m \ge 1,\, x \in V_{(i)} \right\}$$

where

$$w_x(\{n(e)\}_{e \in E_x \setminus E'}) = \frac{\sum_{e \in E_x \setminus E'} n(e)^\alpha}{|E_x \cap E'| + \sum_{e \in E_x \setminus E'} n(e)^\alpha}.$$

Note that K_i^* is not the same as the event that no edge in E' is reinforced from firings at vertices in $V_{(i)}$, as the occurrence of the latter would depend on possible reinforcements of E' from elsewhere. Nevertheless, letting $\mathcal{N}_{E'} := \{E' \subset \mathcal{N}\}$ we have that $\mathcal{N}_{E'} = \cap_{i \le k'} K_i^*$ since while each edge in E' remains unreinforced, the processes $((\bar{N}_t^{(i)}(e))_{e \in E^*_{(i)}})_{i \le k'}$ and $(N_t(e))_{e \in E}$ are identical. They can differ only after some edge in E' is reinforced. For $(A_i)_{i \le k'}$ arbitrary (measurable), but fixed let $B_i^* = \{(\bar{N}_t^{(i)}(e))_{e \in E_i, t \ge 0} \in A_i\}$ and $B_i = \{((N_t(e))_{e \in E_i, t \ge 0} \in A_i\}$. On the event $\mathcal{N}_{E'} = \cap_{i \le k'} K_i^*$ we have that B_i^* occurs if and only if B_i occurs. Hence

$$\mathbb{P}\big(\mathcal{N}_{E'} \cap \cap_{i \le k'} B_i\big) = \mathbb{P}\big(\cap_{i \le k'} \left(B_i^* \cap K_i^*\right)\big).$$

The events $(B_i^* \cap K_i^*)_{i \le k'}$ are independent, as are the events $(K_i^*)_{i \le k'}$, since they depend on independent Poisson processes and uniform random variables. Therefore,

$$\mathbb{P}\left(\cap_{i \le k'} B_i \middle| \mathcal{N}_{E'}\right) = \frac{\mathbb{P}(\cap_{i \le k'}(B_i^* \cap K_i^*))}{\mathbb{P}(\cap_{i \le k'} K_i^*)} = \prod_{i \le k'} \frac{\mathbb{P}(B_i^* \cap K_i^*)}{\mathbb{P}(K_i^*)} = \prod_{i \le k'} \mathbb{P}(B_i^* | K_i^*).$$

Similarly,

$$\mathbb{P}(B_j | \mathcal{N}_{E'}) = \frac{\mathbb{P}(B_j^*, \cap_{i \le k'} K_i^*)}{\mathbb{P}(\cap_{i \le k'} K_i^*)} = \frac{\mathbb{P}(B_j^* \cap K_j^*) \prod_{i \le k', i \ne j} \mathbb{P}(K_i^*)}{\prod_{i' \le k'} \mathbb{P}(K_{i'}^*)} = \mathbb{P}(B_j^* | K_j^*). \tag{6}$$

Therefore,

$$\mathbb{P}\left(\cap_{i \le k'} B_i \middle| \mathcal{N}_{E'}\right) = \prod_{i \le k'} \mathbb{P}(B_i | \mathcal{N}_{E'})$$

which proves the first claim.

The second claim follows immediately from our construction due to the equality of the Poisson processes, equality of $((\bar{N}_t^{(i)}(e))_{e \in E^*_{(j)}})_{j \le k'}$ and $(N_t(e))_{e \in E}$ on the event $\mathcal{N}_{E'} = \cap_{i \le k'} K_i^*$ (and the already proved independence). ■

Now, we have collected all ingredients to prove Proposition 4.

Proof of Proposition 4. We define the two processes on the same probability space by using Remark (3). Let $\boldsymbol{n} \equiv 1$ and $\boldsymbol{n}'$ be as in the remark and let $\boldsymbol{N}' := (N_t^{\boldsymbol{n}'}(e))_{e \in E \setminus E', t \ge 0}$. Then, as in the remark, $\boldsymbol{N}'$ is a WARM process on G' with the same firing process $\boldsymbol{M}$ and with the same partial ordering of edges (as in the proof of Proposition 3). Thus, on the event $\{E' \subset \mathcal{N}\}$, the two processes evolve identically. That is, $\boldsymbol{N'} = (N_t(e))_{e\in E\setminus E', t\geq 0}$ on this event.

For measurable A, let $B = \{(\boldsymbol{M}, (N_t(e))_{e\in E\setminus E', t\geq 0}) \in A\}$ and similarly let $B' = \{(\boldsymbol{M}, (N'_t(e))_{e\in E\setminus E', t\geq 0}) \in A'\}$. Then,

$$\mathbb{P}(B, E' \subset \mathcal{N}) = \mathbb{P}(B', E' \subset \mathcal{N}). \tag{7}$$

Therefore if $\mathbb{P}(B|E' \subset \mathcal{N}) > 0$ then $\mathbb{P}(B') > 0$.

Observe that B' is independent of $(V_n(x))_{x\in V, n\geq 0}$ since by construction $S_t^{\boldsymbol{n'}}(x) = 1$ for every t, x almost surely by our choice of $\boldsymbol{n'}$. Thus, if $\mathbb{P}(B') > 0$ then by Proposition 2 we have

$$\mathbb{P}(E' \subset \mathcal{N}|B') \geq \prod_{v: E_v \cap E' \neq \varnothing} \prod_{m=1}^{\infty} \frac{(m/\partial_v)^\alpha}{|E'| + (m/\partial_v)^\alpha} > 0. \tag{8}$$

Hence, $\mathbb{P}(E' \subset \mathcal{N}|B') > 0$ and therefore $\mathbb{P}(B', E' \subset \mathcal{N}) > 0$. From (7), this proves that $\mathbb{P}(B, E' \subset \mathcal{N}) > 0$ and therefore that $\mathbb{P}(B|E' \subset \mathcal{N}) > 0$ as required. ■

To prove Proposition 5, we construct a coupling $(\boldsymbol{N}, \boldsymbol{N}^*)$ such that $\boldsymbol{N}^*$ has the same distribution as $\boldsymbol{N}$ and we have a lower bound on the probability that $\boldsymbol{N}^*$ never reinforces edges from E'.

Proof of Proposition 5. It is trivial that if $\mathbb{P}(R, E' \subset \mathcal{N}) > 0$ then $\mathbb{P}(R, E' \subset \mathcal{E}_\infty^c) > 0$, so we must prove the converse. Let $V_{E'}$ denote the set of vertices incident to E'.

Suppose that $\mathbb{P}(R, E' \subset \mathcal{E}_\infty^c) > 0$. Then, there exist $t_0 \in \mathbb{N}$ and $k_0 \in \mathbb{Z}_+$ such that with positive probability all of the following occur:

- R occurs,
- $D_0 = \{$no edge in E' is reinforced after time $t_0\}$,
- $J_0 = \{$before time t_0 there are exactly k_0 firings at vertices in $V_{E'}\}$.

We enrich our existing probability space with a family $\boldsymbol{W} = (W_n(x))_{n\in\mathbb{N}, x\in V_{E'}}$ of i.i.d. Bernoulli$(1/2)$-distributed random variables that are independent of $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{M}$ and a Poisson point process $\boldsymbol{M}''_{V_{E'}}$ on $V_{E'} \times [0, t_0]$ with rates $(\lambda_x)_{x\in V_{E'}}$ (i.e. the same as the rates in $\boldsymbol{M}$ at these sites) that is independent of all of these variables. Let $\boldsymbol{M}_{V_{E'}}[0, t_0]$ denote the restriction of $\boldsymbol{M}$ to sites in $V_{E'}$ and times in $[0, t_0]$.

Then,

$$\boldsymbol{M}' = \boldsymbol{M} \cup \boldsymbol{M}''_{V_{E'}}$$

is a Poisson point process with rate $2\lambda_x$ at $x \in V_{E'}$ during the interval $[0, t_0]$ and rate λ_y at $y \in V$ otherwise.

We define a new process $\boldsymbol{N}'$ from the Poisson point process $\boldsymbol{M}'$ exactly as for the process $\boldsymbol{N}$ defined from $\boldsymbol{M}$, using the same $\boldsymbol{U}, \boldsymbol{V}$ and ordering, and with $\boldsymbol{N}'_0(e) = \boldsymbol{N}_0(e)$ for each $e \in E$.

We also define a new process $\boldsymbol{N}^*$ from the Poisson point process $\boldsymbol{M}'$ as follows. Set $\boldsymbol{N}^*_0(e) = \boldsymbol{N}_0(e)$ for each $e \in E$. At the mth firing (in the process $\boldsymbol{M}'$) at x at some time t, if *either* $[x \in V_{E'}$ and $t \geq t_0]$, or $[x \notin V_{E'}$ and $t \geq 0]$, then we choose edge $\hat{e}^*$ according to

$$\begin{aligned}\mathbb{1}_{\{\hat{e}^*=e\}} = {} & \mathbb{1}_{\{e \in E_x \cap E^-\}} \mathbb{1}_{\left\{V_m(x) \leq S^*_{t-}(x)\right\}} \mathbb{1}_{\{U_m(x) \in (Q^*_t(x,e), Q^*_t(x,e)+R^*_t(e)]\}} \\ & + \mathbb{1}_{\{e \in E_x \cap E'\}} \mathbb{1}_{\left\{V_m(x) > S^*_{t-}(x)\right\}} \mathbb{1}_{\left\{U_m(x) \in (Q'^*_t(x,e), Q'^*_t(x,e)+R'^*_t(e)]\right\}},\end{aligned}$$

where the starred quantities are defined as in Section 2 but for the process $\boldsymbol{N}^*$ (using the same $\boldsymbol{U}, \boldsymbol{V}$ variables and ordering of edges). We then set $\boldsymbol{N}^*_t(e) = \boldsymbol{N}^*_{t-}(e) + \mathbb{1}_{\{e=\hat{e}^*\}}$.

If (in the process $\boldsymbol{M}'$) we have $(X_n, T_n) = (x, t)$ for some $x \in V_{E'}$ and $t < t_0$, we don't reinforce any edge if $W_n(x) = 0$, while if $W_n(x) = 1$ and there have been m firings at x in $[0, t]$ then we choose edge $\hat{e}^*$ according to

$$\begin{aligned}\mathbb{1}_{\{\hat{e}^*=e\}} = {} & \mathbb{1}_{\{e \in E_x \cap E^-\}} \mathbb{1}_{\left\{V_m(x) \leq S^*_{t-}(x)\right\}} \mathbb{1}_{\{U_m(x) \in (Q^*_t(x,e), Q^*_t(x,e)+R^*_t(e)]\}} \\ & + \mathbb{1}_{\{e \in E_x \cap E'\}} \mathbb{1}_{\left\{V_m(x) > S^*_{t-}(x)\right\}} \mathbb{1}_{\left\{U_m(x) \in (Q'^*_t(x,e), Q'^*_t(x,e)+R'^*_t(e)]\right\}},\end{aligned}$$

and we set $\boldsymbol{N}^*_t(e) = \boldsymbol{N}^*_{t-}(e) + \mathbb{1}_{\{e=\hat{e}^*\}}$.

Notice that $\boldsymbol{N}^*$ only reinforces an edge at firings in the set

$$\boldsymbol{M}^* = \{(X_n, T_n) \in \boldsymbol{M}' : W_{\tau'_n}(X_n) = 1 \text{ if } X_n \in V_{E'} \text{ and } T_n \leq t_0\},$$

where $\tau'_n = |\{k \leq n : X_k = X_n\}|$ is the number of times that X_n has fired (in the process $\boldsymbol{M}'$) up to time T_n. Clearly, $\boldsymbol{M}^*$ is the process $\boldsymbol{M}$ filtered on $V_{E'} \times [0, t_0]$ by $\boldsymbol{W}$, which is easily seen to be a Poisson point process on $V \times [0, \infty)$ with rates $\boldsymbol{\lambda}_V$. Thus, the processes $(\boldsymbol{M}, \boldsymbol{N})$ and $(\boldsymbol{M}^*, \boldsymbol{N}^*)$ have the same law.

Let K denote the event that: $W_n(X_n) = 0$ for each of the firings $(X_n, T_n) \in \boldsymbol{M}'_{V_{E'}}[0, t_0]$ that result in a choice of an edge in E', and $W_n(X_n) = 1$ for each of the firings $(X_n, T_n) \in \boldsymbol{M}'_{V_{E'}}[0, t_0]$ that result in a choice of an edge in $E \setminus E'$. Let L' denote the number of firings in $\boldsymbol{M}'_{V_{E'}}[0, t_0]$. Then,

$$\begin{aligned}\mathbb{P}(R, D_0, J_0, K) &= \sum_{\ell \geq k_0} \mathbb{P}(R, D_0, J_0, K | L' = \ell)\mathbb{P}(L' = \ell) \\ &= \sum_{\ell \geq k_0} \mathbb{P}(R, D_0, J_0 | L' = \ell)\mathbb{P}(K | L' = \ell)\mathbb{P}(L' = \ell),\end{aligned}$$

where we have used the fact that K is independent of R, D_0, J_0 given $L' = \ell$. Since $\mathbb{P}(K|L' = \ell) = 2^{-\ell}$ we have

$$\mathbb{P}(R, D_0, J_0, K) = \sum_{\ell \geq k_0} 2^{-\ell}\mathbb{P}(R, D_0, J_0, L' = \ell),$$

which is strictly positive since $0 < \mathbb{P}(R, D_0, J_0) = \sum_{\ell \geq k_0} \mathbb{P}(R, D_0, J_0, L' = \ell)$.

Now, observe that on the event $D_0 \cap J_0 \cap K$ we have that $\boldsymbol{N}^*_{E \setminus E'} = \boldsymbol{N}_{E \setminus E'}$ (so in particular on $D_0 \cap J_0 \cap K$, R^* occurs if and only if R^* occurs) and moreover $E' \subset \mathcal{N}^*$. Thus,

$$\mathbb{P}(R^*, E' \subset \mathcal{N}^*) \geq \mathbb{P}(R, D_0, J_0, K) > 0.$$

Now, recall that the processes $(\boldsymbol{M}, \boldsymbol{N})$ and $(\boldsymbol{M}^*, \boldsymbol{N}^*)$ have the same law, so $\mathbb{P}(R, E' \subset \mathcal{N}) = \mathbb{P}(R^*, E' \subset \mathcal{N}^*) > 0$ as claimed. ■

Proof of Proposition 6. First, Proposition 1 gives that

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{E}_+ = E \setminus E') = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{E}^c_\infty = E') = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(E' \subset \mathcal{E}^c_\infty, R),$$

where $R = \{\sup_{t \geq 0} N_t(e) = \infty$ for every $e \in E \setminus E'\}$. Thus, by Propositions (5) and (2), we have

$$\begin{aligned}\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{E}_+ = E \setminus E') > 0 &\iff \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(E' \subset \mathcal{N}, R) > 0 \\ &\iff \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R | E' \subset \mathcal{N}) > 0.\end{aligned}$$

Putting $G' = (V, E \setminus E')$, Proposition 4 gives that

$$\begin{aligned}\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R | E' \subset \mathcal{N}) > 0 &\iff \mathbb{P}_{G',\boldsymbol{\lambda}_V,\alpha}(R) > 0 \\ &\iff \prod_{i \leq k'} \mathbb{P}_{G_i,\boldsymbol{\lambda}_{V_i},\alpha}(R_i) > 0 \\ &\iff \prod_{i \leq k'} \theta_{G_i,\boldsymbol{\lambda}_{V_i}}(\alpha) > 0 \\ &\iff \theta_{G_i,\boldsymbol{\Lambda}_{V_i}}(\alpha) > 0 \text{ for every } i \leq k',\end{aligned}$$

where $(G_i)_{i\leq k'}$, are the disjoint components of G', R_i is the event that every edge in E_i is used infinitely often, and where in the second line we have used the fact that a WARM process is independent on disjoint components. ■

5 Proof of Theorem 1

In this Section we prove Theorem 1, which asserts that when E is finite $(X_e(n))_{e\in E}$ converges almost surely to a random vector *supported on the set of not-linearly-unstable equilibria* of a particular deterministic dynamical system. In order to introduce this system, we generalise the model to the general setting considered in [8, 10]. In this setting we have a finite set E of colours of balls and collection of probabilities $\boldsymbol{p} = (p_A)_{A\subset E}$, with $p_\varnothing = 0$. At each step of the process we choose a subset $A \subset E$ (with probability p_A independent of the history of the process) of the colours to compete for one step of a Pólya urn process with parameter α. The dynamics of counts of the edges in a finite, discrete-time WARM process are given by

$$N_e(n+1) = N_e(n) + \eta_{n,e},$$

where $\eta_{n,e}$ are random variables, taking values 0 or 1, and indicating the random reinforced edge at time n. We have

$$\mathbb{P}(\eta_{n,e} = 1|\mathcal{F}_n) = \sum_{A\ni e} p_A \frac{(N_e(n))^\alpha}{\sum_{j\in A}(N_j(n))^\alpha},$$

where $\mathcal{F}_n = \sigma((N_e(k))_{e\in E}, k \leq n)$. Setting $\rho_{n,e} := \eta_{n,e} - X_e(n)$, the proportions $X_e(n) = \frac{N_e(n)}{n}$ satisfy

$$X_e(n+1) = X_e(n) + \frac{1}{n+1}(\eta_{n,e} - X_e(n)) = X_e(n) + \frac{1}{n+1}\rho_{n,e}. \tag{9}$$

Thus,

$$\mathbb{E}[X_e(n+1) - X_e(n)|\mathcal{F}_n] = \frac{1}{n+1}F_e(\vec{X}(n)),$$

where

$$F_e(\vec{x}) = \sum_{A\ni e} p_A \frac{x_e^\alpha}{\sum_{j\in A} x_j^\alpha} - x_e, \text{ for } e \in E. \tag{10}$$

The solutions $\vec{v}$ of this system of equations are *equilibria*. Let $\mathcal{J}$ denote the Jacobian of $(F_e)_{e\in E}$. An equilibrium $\vec{v}$ is *linearly unstable* if at least one eigenvalue of $\mathcal{J}(\vec{v})$ is strictly positive.

It is known for instance from [8, Theorem 1] that all accumulation points of $\boldsymbol{X}$ are equilibria of the above deterministic system, and the argument of [3, Section 4] shows that if $\mathcal{U}$ is the set of linearly unstable equilibria, then $\mathbb{P}(\lim_{n\to\infty}(X_e(n))_{e\in E} \in \mathcal{U}) = 0$. Therefore to prove Theorem 1 it remains to show that $\mathbb{P}(\lim_{n\to\infty}(X_e(n))_{e\in E}$ exists$) = 1$.

To this end, we will prove the following result, by adapting arguments of [16], in which

$$\underline{X}_E = \min_{e\in E} \inf_{n\geq 0} X_e(n),$$

denotes the smallest rescaled weight over all edges and times and

$$D = \big\{ \liminf_{n\to\infty} \boldsymbol{X}(n) \neq \limsup_{n\to\infty} \boldsymbol{X}(n) \big\}$$

denotes the event that $\boldsymbol{X}(n)$ does not converge.

Proposition 8. *For any finite WARM, with $\alpha > 0$ and $\boldsymbol{\lambda}_V > 0$*

$$\mathbb{P}_{G,\alpha,\boldsymbol{\lambda}_V}(\underline{X}_E > 0, D) = 0.$$

Proof of Theorem 1 assuming Proposition 8, As explained in Section 1, it suffices to consider the discrete-time dynamics. We first prove by induction on the number of edges in G that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(D) = 0$. If G has only one edge then the result is trivial. Let G be a graph containing n edges. By Proposition 8 it suffices to show that $\mathbb{P}_{G,\alpha,\boldsymbol{\lambda}_V}(\underline{X}_E = 0, D) = 0$. Now,

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\underline{X}_E = 0, D) \leq \sum_{e\in E} \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}\Big(\inf_{n\geq 0} X_e(n) = 0, D\Big),$$

so it is sufficient to show that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\inf_{n\geq 0} X_e(n) = 0, D) = 0$. Suppose instead that this probability is strictly positive. Then, by Proposition 1, $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(e \in \mathcal{E}_\infty^c, D) > 0$. In particular, $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(e \in \mathcal{E}_\infty^c) > 0$ and this implies that e must be removable, otherwise $N_e(n)$ would grow asymptotically linearly in time almost surely. Since $X_e(n)$ converges to 0 on the event $\{e \in \mathcal{E}_\infty^c\}$, we must have that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(e \in \mathcal{E}_\infty^c, D'_e) > 0$ where D'_e is the event that $X_u(n)$ does not converge for some $u \neq e$. Now, Propositions 5 and 4 imply that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(e \in \mathcal{N}, D'_e) > 0$ and $\mathbb{P}_{G\setminus\{e\},\boldsymbol{\lambda}_V,\alpha}(D) > 0$. But this violates the induction hypothesis. We therefore conclude that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\inf_{n\geq 0} X_e(n) = 0, D) = 0$ and hence $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(D) = 0$ as claimed. ■

The idea of proof for Proposition 8 is to first implement a change of variables transforming the vector field F into an anti-gradient, and then to apply the main result of [16].

Proof of Proposition 8. An easy calculation gives

$$F_e(\vec{x}) = -x_e \frac{\partial L}{\partial x_e},$$

where

$$L(\vec{x}) = \sum_{e\in E} x_e - \frac{1}{\alpha}\sum_{A\subset E} p_A \log\Big(\sum_{e\in A} x_e^\alpha\Big). \tag{11}$$

Hence, the vector field F becomes anti-gradient after the change of variables

$$\Phi(\vec{x}) = \vec{y}, \quad y_e = 2\sqrt{x_e}.$$

That is, the image $y(t) = \Phi(x(t))$ under Φ of a trajectory $x(t)$ of a vector field $\dot{x} = F(x)$ satisfies

$$\begin{aligned} \dot{y}_e &= \frac{1}{\sqrt{x_e}} \cdot \dot{x}_e = \frac{1}{\sqrt{x_e}} \cdot (-x_e \frac{\partial L}{\partial x_e}) \\ &= -\sqrt{x_e}\,\frac{\partial L\circ\Phi^{-1}}{\partial y_e}\frac{\partial y_e}{\partial x_e} = -\sqrt{x_e}\cdot\frac{1}{\sqrt{x_e}}\cdot\frac{\partial L\circ\Phi^{-1}}{\partial y_e} = -\frac{\partial \mathcal{L}}{\partial y_e}, \end{aligned} \tag{12}$$

where $\mathcal{L} := L\circ\Phi^{-1}$. Let $\mathcal{D} := \{(y_e)_{e\in E} : \min_{e\in E} y_e \geq 2\sqrt{\varepsilon}\}$.

Now, we apply this change of variables to the WARM process. That is, we consider the random increments

$$r_{n,e} = Y_e(n+1) - Y_e(n) \tag{13}$$

of the random variables $Y_e := 2\sqrt{X_e}$. If we can show that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\underline{X}_E > \varepsilon, D) = 0$ for every $\varepsilon > 0$, then $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\underline{X}_E > 0, D) = 0$.

Hence, let $\varepsilon > 0$ be arbitrary. Then, $X_e(n) > \varepsilon$ for every $e \in E$ and every $n \geq 0$, so that

$$X_e(n+1) - X_e(n) = \frac{1}{n+1}\rho_{n,e} = O(\frac{1}{n+1}).$$

Applying a map Φ that is smooth in the domain $\{\boldsymbol{x} : \min_{e\in E} x_e > \varepsilon\}$, we obtain

$$r_{n,e} = \frac{\partial y_e}{\partial x_e}|_{X_e(n)} \cdot \frac{\rho_{n,e}}{n+1} + O\Big(\frac{\rho_{n,e}^2}{(n+1)^2}\Big).$$

Now, taking the conditional expectation w.r.t. $\mathcal{F}_n$, we get

$$\begin{aligned} \mathbb{E}[r_{n,e} \mid \mathcal{F}_n] &= -\frac{1}{n+1}\frac{\partial y_e}{\partial x_e}|_{X_e(n)}\mathbb{E}[\rho_{n,e} \mid \mathcal{F}_n] + O(\frac{1}{n^2}) \\ &= -\frac{1}{n+1}\frac{\partial \mathcal{L}}{\partial y_e}|_{Y_e(n)} + O(\frac{1}{n^2}), \end{aligned} \tag{14}$$

where we have used that $\mathbb{E}[\rho_{n,e} \mid \mathcal{F}_n] = F_e(\vec{X}(n))$ and (12).

Hence, the process (13) can be represented as

$$\begin{aligned} Y_e(n+1) &= Y_e(n) + \mathbb{E}[r_{n,e} \mid \mathcal{F}_n] + (r_{n,e} - \mathbb{E}[r_{n,e} \mid \mathcal{F}_n]) \\ &= Y_e(n) - \frac{1}{n+1}\Big(\frac{\partial \mathcal{L}}{\partial y_e}|_{Y_e(n)} + \xi_{n,e}\Big), \end{aligned}$$

where the "random term"

$$\xi_{n,e} := (n+1)\Big(\big(\mathbb{E}[r_{n,e} \mid \mathcal{F}_n] - r_{n,e}\big) + \Big(-\frac{1}{n+1}\frac{\partial \mathcal{L}}{\partial y_e} - \mathbb{E}[r_{n,e} \mid \mathcal{F}_n]\Big)\Big). \quad (15)$$

decomposes into two drift components: the difference of r_n with its (conditional) expectation and the difference between the conditional expectation of r_n and the $\mathcal{L}$-antigradient flow step.

In the right hand side of (15), the former summand is of zero expectation, while the latter one is estimated via (14). Thus, $\mathbb{E}[\xi_{n,e}] = O(\frac{1}{n})$.

This is exactly the setting considered in [16], which is devoted to the study of the stochastic gradient flow: Eq. (1) therein defines a sequence of random variables taking values in $\mathbb{R}^k$

$$\theta_{n+1} = \theta_n - \alpha_n(\nabla f(\theta_n) + \xi_n).$$

Then, [16, Theorem 2.1] states that under several assumptions the limit of such a sequence θ_n exists almost surely. Let us check these assumptions. The first one, Assumption 2.1, is almost immediate: up to a constant shift of the time index, in our case, we have $\alpha_n = \frac{1}{n}$, and hence

$$\lim_{n\to\infty} \alpha_n = 0, \quad \sum_{n\geq 1} \alpha_n = \infty.$$

Next, Assumption 2.3 of [16] in any compact domain is implied by the analyticity of the function $\mathcal{L}$ (see the discussion on the two last paragraphs of [16, p. 1717]) as it is exactly the Lojasiewicz inequality. The function $\mathcal{L}$ is analytic in the domain $\mathcal{D}$, and the main theorem in [16] in fact (by the same arguments) also implies the almost sure convergence on the set of trajectories that stay in a domain where the assumptions are satisfied.

The most technical assumption is Assumption 2.2. Let $\gamma_n := \sum_{i<n} \alpha_i$, and define the moment of time $a(n,s)$ as

$$a(n,s) := \max\{k \geq n : \gamma_k - \gamma_n \leq s\}. \quad (16)$$

Then, Assumption 2.2 requires that for some $r > 1$ for any s the random variable

$$\xi := \limsup_{n\to\infty} \max_{n\le k\le a(n,s)} \left\| \sum_{n\le i\le k} \alpha_i \gamma_i^r \xi_i \right\| \tag{17}$$

is almost surely finite on the set of trajectories staying in $\mathcal{D}$. Very roughly speaking, it is this assumption where one controls both the created variance and the drift, coming from the random variables ξ_k.

For our case we have $\alpha_n = \frac{1}{n}$, hence

$$\gamma_k = \log k + C + o(1), \quad a(n,s) = (e^s + o(1)) \cdot n.$$

As larger values of s imply larger values of $a(n,s)$, and the right hand side of (17) is increasing in $a(n,s)$, it suffices to check the finiteness for $2n$ taken instead of $a(n,s)$ (as it also implies the finiteness for $4n$, $8n$, etc.).

Let us check this assumption; in fact, we will show that any r (in particular, $r = 2$) will do. We start by decomposing

$$\xi_n = q_n + \bar{\xi}_n,$$

where $q_n := \mathbb{E}[\xi_n \mid \mathcal{F}_n]$. Recall that we have $q_n = O(\frac{1}{n})$. Also, as $|\rho_{t,e}| \le \frac{1}{t+1}$, the random variables ξ_n's are uniformly bounded.

Then, the series $\sum_{n\ge1} |\alpha_n \gamma_n^r q_n|$ converges since the bound for q_n implies that its terms are bounded by $O(\frac{\log^r n}{n^2})$. On the other hand, the sequence of intermediate sums

$$S_n := \sum_{n\ge1} \alpha_n \gamma_n^r \bar{\xi}_n$$

is a (coordinatewise) martingale, and we have

$$\mathbb{E}[(\alpha_n \gamma_n^r \bar{\xi}_{n,e})^2 | \mathcal{F}_{n-1}] = O(\frac{\log^{2r} n}{n^2})$$

with a uniform constant. As the series $\sum_{n\ge1} \frac{\log^{2r} n}{n^2}$ converges, the martingale $S_{n,e}$ has a uniformly bounded second moment and hence converges almost surely. Thus, the same applies to the series

$$\sum_{n\ge1} \alpha_n \gamma_n^r \xi_{n,e} = \sum_{n\ge1} \alpha_n \gamma_n^r q_{n,e} + \sum_{n\ge1} \alpha_n \gamma_n^r \bar{\xi}_{n,e}.$$

Hence, almost surely the series

$$\sum_{n\ge1} \alpha_n \gamma_n^r \xi_n$$

converges (as it converges coordinatewise), provided that along the iterations we stay in the domain $\mathcal{D}$. In particular, the pairwise differences of its intermediate sums, that are of the form

$$\sum_{n \le i \le k} \alpha_i \gamma_i^r \xi_i,$$

are almost surely bounded uniformly in n and k. In particular, we get the desired

$$\limsup_{n\to\infty} \max_{n \le k < a(n,t)} \left\| \sum_{n \le i \le k} \alpha_i \gamma_i^r \xi_i \right\| < \infty.$$

■

6 Proof of Theorems 2 and 3

Using the auxiliary results from Section 3, we now complete the proof of Theorem 3.

Proof of Theorem 3. We claim that

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \mathcal{A}_F) > 0 \Leftrightarrow \partial F \text{ is removable and } \mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(R, \mathcal{E}_+ = F) > 0. \tag{18}$$

Once (18) is established, the first claim follows by taking $R = \Omega$. Moreover, given the the first claim holds, if $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{A}_F) > 0$ then also $\mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(\mathcal{E}_+ = F) > 0$ and vice versa so we can replace (18) with

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R | \mathcal{A}_F) > 0 \Leftrightarrow \partial F \text{ is removable and } \mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(R | \mathcal{E}_+ = F) > 0, \tag{19}$$

which proves the second claim.

To prove (18), note that by definition of $\mathcal{A}_F$ and Proposition 1 and Theorem 1, for any $R \in \sigma(N_t(e) : t \ge 0, e \in F)$ we have

$$\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \mathcal{A}_F) = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \partial F \subset \mathcal{E}_+^c, F \subset \mathcal{E}_+) = \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \partial F \subset \mathcal{E}_\infty^c, F \subset \mathcal{E}_+).$$

Moreover, by Proposition 5 the right-hand side is positive if and only if $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \partial F \subset \mathcal{N}, F \subset \mathcal{E}_+) > 0$. Hence, applying Proposition 2 and Proposition 4 with $G' = (V, E \setminus \partial F)$, we arrive at

$$\begin{aligned}\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, \mathcal{A}_F) > 0 &\Leftrightarrow \partial F \text{ is removable and } \mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(R, F \subset \mathcal{E}_+ | \partial F \subset \mathcal{N}) > 0 \\ &\Leftrightarrow \partial F \text{ is removable and } \mathbb{P}_{G',\boldsymbol{\lambda}_V,\alpha}(R, F \subset \mathcal{E}_+) > 0,\end{aligned}$$

But since F is a connected component of G' and the WARM process behaves independently on different connected components we have that

$$\mathbb{P}_{G',\boldsymbol{\lambda}_V,\alpha}(R, F \subset \mathcal{E}_+) = \mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(R, \mathcal{E}_+ = F),$$

which gives (18). ■

We now use Theorem 3 to prove Theorem 2.

Proof of Theorem 2. Let $\alpha > 2$ (resp. $\alpha > 25$). Suppose that with positive probability $\mathcal{E}_+$ contains a finite connected component for which the first claim (resp. second claim) of the theorem fails. Since the set consisting of the finite connected subsets of G is countable (here we are using the fact that G is countable), there must exist a finite connected subgraph $F \subset G$ for which $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{A}_F) > 0$ but such that F is not a tree (resp. has diameter larger than 3). By Theorem 3, we have $\mathbb{P}_{F,\boldsymbol{\lambda}_{V_F},\alpha}(\mathcal{E}_+ = F) > 0$. By Theorem 1 the limits exist a.s. and there exists $\boldsymbol{x} \in \mathcal{S}$ such that $\sigma(\mathcal{S}) = F$. By [10, Corollary 1] F must be a tree (resp. [10, Theorem 2(a)], F must have diameter at most 3), which gives a contradiction. ■

7 Proof of Theorem 4 and Corollary 1

Henceforth we assume that Condition 2 holds. The proof of Theorem 4 is based on the following ancillary results.

Proposition 9. *For $M \geq 4$ and $\alpha > 2$, all fully supported equilibria on $[0, M]$ are linearly unstable.*

Proposition 10.

(i) There is no equilibrium supported on $[0, 3]$ for $\alpha \in (1, \alpha^*)$.

(ii) There is a linearly stable equilibrium supported on $[0, 3]$ for $\alpha > \alpha^*$.

Proposition 10(ii) is proved in [8, proof of Theorem 8] (take $r = 1$ in that result), so we will only need to prove (i).

Finally, for α close to 1, large connected components appear, see e.g. Figure 4.

Proposition 11. *Let $k \geq 1$ be arbitrary. Then, there exists $\alpha_{2k} > 1$ such that for $\alpha \in (1, \alpha_{2k})$ the graph $G = [0, 2k]$ admits a fully-supported linearly stable equilibrium with weights $\vec{x}(k; \alpha) = (x_i(k; \alpha))_{i \in 0,\ldots,2k-1}$ such that*

$$\lim_{\alpha \downarrow 1} \vec{x}(k; \alpha) = \frac{1}{k(k+1)}(k, 1, k-1, 2, \ldots, k-1, 1, k), \tag{20}$$

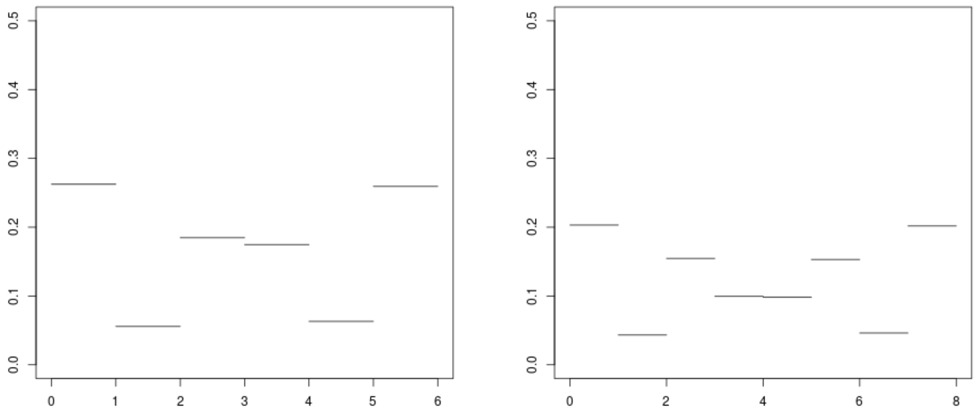


Figure 4: Approximations of the stable equilibria in Proposition 11 on the intervals $[0,6]$ and $[0,8]$ respectively when $\alpha = 1.1$. The pictures in each case are based on simulating the process for a large number of steps, followed by deterministic dynamics to speed up convergence.

The proofs of Propositions 9, 10, and 11 will be deferred to the appendix. We now explain why they imply Theorem 4.

Proof of Theorem 4. To prove the claim (0) first note that the set of edges $E' = \{(2i, 2i+1) : i \in \mathbb{Z}\}$ is removable. Let $A_i = \{(2i, 2i+1) \in \mathcal{N}\}$ denote the event that the edge $(2i, 2i+1)$ is never used. The events A_i are dependent, however given the occurrence or otherwise of each A_j for $j \neq i$, A_i has probability at least

$$p_\alpha := \left(\prod_{k \in \mathbb{N}} \frac{k^\alpha}{k^\alpha + 1} \right)^2 > 0$$

of occurring. The lower bound is the probability that the middle edge is never used *in the graph* $[0,3]$ *with firing rates equal to 0 at the extremities.* It arises by considering a pair of sequences of uniform random variables used to choose edges at firing times at the sites $2i$ and $2i+1$ respectively, similarly to the construction in Section 2.

It is elementary that $\theta_{L_i}(\alpha) = 1$ for line graphs L_1, L_2 containing 1 and 2 edges respectively (since in both cases each edge is incident to a leaf). Claim (i) of the Theorem now follows from Theorem 3 since both $\{(-1,0),(1,2)\}$ and $\{(-1,0),(2,3)\}$ are removable sets of edges. Claims (ii), (iii) and (iv) follow similarly from Theorems 1 and 3 combined with Propositions 9, 10, and 11. ■

The proof of Corollary 1 is based on the following key result on the instability of equilibria supported on even cycles. This is a generalisation of [8, Proposition 2], i.e. that the uniform distribution on any even cycle is linearly unstable for all $\alpha > 1$.

Lemma 2. *Let G be a finite graph containing an even cycle, $\alpha > 1$ and $\delta = 1$. Then, any fully supported equilibrium on G is linearly unstable.*

Proof. Recall first from [8] (see also (11)) that in the appropriately chosen coordinates, our flow is an anti-gradient one, associated with the function

$$L(\vec{x}) = \sum_{i \leq n} x_i - \frac{1}{\alpha \cdot |V|} \sum_{v \in V} \log\Big(\sum_{e_i \sim v} x_i^\alpha\Big). \tag{21}$$

Hence, if for an equilibrium point (that is, a critical point of L) there exists a direction for which the second derivative of L is strictly negative, this equilibrium is linearly unstable. We will find such a direction by finding a parametric curve $\vec{x} = \vec{x}(s)$, such that $\vec{x}(0)$ is our equilibrium and the second derivative $\frac{d^2}{ds^2} L(\vec{x}(s))$ is strictly negative.

To construct such a curve, let us first pass to the coordinates $y_i = x_i^\alpha$. Then, $x_i = y_i^{1/\alpha}$ is a concave function of y_i. Now, assume that the colours $i = 1, \dots, 2m$ form the even cycle from the assumptions of the lemma. Then, given an initial point x and the corresponding powers y, consider a straight segment in the y coordinates: let

$$\tilde{y}_j(s) = \begin{cases} y_j + (-1)^j \cdot s & j \leq 2m, \\ y_j & \text{otherwise.} \end{cases}$$

Then, for any vertex on the even cycle the corresponding sum $y_j + y_{j+1}$ does not change along this segment. Hence, the logarithmic part of the function L stays unchanged on this curve. At the same time, each x_i is a concave function of s, and hence their sum also is. Thus, the restriction of L to the curve $\tilde{x}_j(s) = \tilde{y}_j(s)^{1/\alpha}$ is strictly concave, what concludes the proof. ■

Proof of Corollary 1. Suppose that with positive probability there exists a finite component F of $\mathcal{E}_+$ containing an even cycle. Since the collection of finite subsets of edges is countable there must exist a finite F containing an even cycle such that $\mathbb{P}_{G,\boldsymbol{\lambda}_V,\alpha}(\mathcal{A}_F) > 0$. Thus by Theorem 3, $\theta_{F,\boldsymbol{\lambda}_{V_F},\alpha} > 0$. By Theorem 1 there must be a fully supported equilibrium for the WARM on F that is not linearly unstable. This contradicts Lemma 2, and therefore proves the first claim of the Corollary.

For the second claim ▲, observe that in the triangular lattice for each $m \geq 1$ there is an odd cycle C_n of length $n = 2m + 1$ containing the origin, such that $(\partial C_n)^c$ consists of two disjoint components: a single infinite connected component, and C_n (i.e. there are no vertices in the "interior" of C_n). Thus, ∂C_n is removable. Moreover, [8, Proposition 2(iv)] implies that if α is sufficiently close to 1, then the uniform distribution on C_n is linearly stable. The claim ▲ now follows from Theorems 1, 3, and Proposition 7.

The proof of the final claim, ★ is similar, using Theorem 4. ■

A Appendix

Here we prove Propositions 9, 10, and 11, with the first of these being the most difficult.

Given an equilibrium state $(a_e)_{e\in E}$ on some finite graph, for each edge e and each vertex v incident to it denote by $q_{e,v}$ the proportion of the firings at this vertex that (on average) goes to e:

$$q_{e,v} := \frac{a_e^\alpha}{\sum_{f\sim v} a_f^\alpha}. \tag{22}$$

The following necessary condition for an equilibrium to be not linearly unstable was shown in [10] (from taking second partial derivatives of the function L).

Lemma 3 ([10, Eq. (18)]). *Let $G = (V, E)$ be a finite graph and $(a_e)_{e\in E}$ be an equilibrium that is not linearly unstable. Then for any e such that $a_e > 0$ we have*

$$\sum_{v\sim e} \lambda_v \cdot q_{e,v}(1 - \alpha(1 - q_{e,v})) \geq 0. \tag{23}$$

In particular, for the case of the line graph $[0, M]$ (with vertices $V = \{0, 1, \dots, M\}$ and edges labelled 1 to M according to their right endpoint) and $\boldsymbol{\lambda}_V \equiv 1$, this immediately implies the following.

Corollary 2. *Let $M \geq 1$, and $\alpha > 1$. Let $(a_i)_{i\in[M]}$ be a fully supported equilibrium on the line graph $[0, M]$ (with $\boldsymbol{\lambda}_V \equiv 1$) that is not linearly unstable. Then for each $i = 1, \dots, M$,*

$$q_{i,i-1}(1 - \alpha(1 - q_{i,i-1})) + q_{i,i}(1 - \alpha(1 - q_{i,i})) \geq 0. \tag{24}$$

Another consequence of Lemma 3 is that for $\alpha > 1$, for any equilibrium that is not linearly unstable, every surviving edge e takes a share $q_{e,v} \geq 1 - \frac{1}{\alpha}$ from at least one of the adjacent vertices v; for $\alpha > 2$, it is more than half of these firings, so following the teminology of [10], we call it the *champion* at v. Indeed, the product

$$\varphi(q) := q(1 - \alpha(1 - q)),$$

that is under the summation sign in (24), is positive on the interval $[0, 1]$ only for $q > 1 - \frac{1}{\alpha}$. For the line graph, this also implies the following.

Corollary 3. *Let $\alpha > 2$ and $M \geq 2$. For any fully supported equilibrium $(a_i)_{i \in [M]}$ on the line graph $[0, M]$ (with $\boldsymbol{\lambda}_V \equiv 1$) that is not linearly unstable, there exists $m \in \{1, \ldots, M - 1\}$ such that*

$$a_1 < \cdots < a_m, \quad \text{and} \quad a_{m+1} > a_{m+2} > \cdots > a_M,$$

where if $m = 1$ (resp. $m = M - 1$) the left (resp. right) hand chain of inequalities makes no assertion.

Proof. As $\alpha > 2$, the only way for an inner edge to be selected more than half of the time from a particular adjacent vertex is to have a weight strictly greater than the other edge adjacent to it. Thus the weights decrease monotonically as one goes away from an edge of maximal weight $\max_i a_i$. ∎

Proof of Proposition 9 (assuming Lemmas 4 and 5 below). We prove the result by contradiction. Let $M \geq 4$ and $\alpha > 2$, and suppose that $(a_i)_{i \in [M]}$ is a fully supported equilibrium on $[0, M]$ that is not linearly unstable. For convenience we will assume that the a_i have not been normalised, i.e. $\sum_i a_i = \sum_v \lambda_v$. Consider the next-to-the-boundary edges, with weights a_2 and a_{M-1}. Upon reflecting if necessary, we can assume that $a_2 \leq a_{M-1}$, and by Corollary 3 we have that $a_1 < a_2 \leq a_3$. We will then see that the equilibrium conditions are incompatible (for $\boldsymbol{\lambda}_V \equiv 1$) with the condition (24) for a particular edge $i = 2$. This will provide us with the desired contradiction.

Namely, let

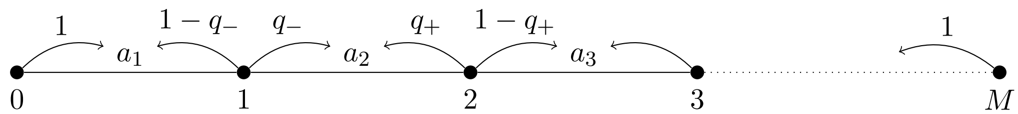
$$q_- = q_{2,1} = \frac{a_2^\alpha}{a_1^\alpha + a_2^\alpha}, \quad q_+ = q_{2,2} = \frac{a_2^\alpha}{a_2^\alpha + a_3^\alpha} \tag{25}$$

be the proportions of firings that this edge is getting from the adjacent vertices. Since $a_2 \leq a_3$ we have $q_+ \leq 1/2$, and therefore as discussed after

Corollary 2 we must have $q_- \geq 1 - 1/\alpha$. Moreover, recalling that the a_i are not normalised and $\lambda_v = 1$ for all v, we have

$$a_1 = 1 + (1 - q_-), \quad a_2 = q_- + q_+. \tag{26}$$

Let us now split the consideration into two cases, "small" $\alpha \leq 7$ and "large"



$\alpha > 7$.

When $\alpha \leq 7$ we will show that for any $q_+ \leq \frac{1}{2}$ there is no solution to the system of equations on q_-, a_1, a_2 such that $a_2 > a_1$ (that is required to ensure the condition (24)). Roughly speaking, the edges $[0,1]$ and $[1,2]$ get from their outer vertices 0 and 2 proportions $q_{1,0} = 1$ and $q_+ \leq \frac{1}{2}$ respectively, while competing via their common vertex 1. It turns out that the power $\alpha \leq 7$ is not sufficiently high for there to exist an equilibrium with $a_2 > a_1$; this is done in Lemma 4 below.

To handle the case $\alpha > 7$, we note that since $q_- \leq 1$, (24) implies that

$$\varphi(q_+) \geq -1,$$

and in turn it implies an upper bound for αq_+ by a constant $C := 3 - \sqrt{2}$; see Lemma 5 below. Note that once such a bound is established, the rest is easy. Indeed, we have an upper bound

$$a_2 \leq 1 + q_+ \leq 1 + \frac{C}{\alpha};$$

thus

$$(a_2/a_1)^\alpha \leq (1 + \frac{C}{\alpha})^\alpha < e^C < 4.9.$$

This implies that the proportion

$$1 - q_- = q_{1,1} = \frac{1}{1 + (a_2/a_1)^\alpha} > \frac{1}{5.9}.$$

which contradicts $q_- > 1 - \frac{1}{\alpha}$ since $\alpha > 7$. ■

Let us prove the two remaining statements:

Lemma 4. *For any $2 < \alpha \leq 7$ there exist no positive a_1, a_2, a_3, q_- such that $a_1 < a_2 \leq a_3$ and that (25) and (26) are satisfied.*

Proof. Assume the contrary. Then from $a_2 \leq a_3$ we have $q_+ \leq \frac{1}{2}$, and hence

$$a_2 = q_- + q_+ \leq q_- + \frac{1}{2}.$$

On the other hand, $a_1 = 1 + (1 - q_-) = 2 - q_-$, and the quotient between the edges' weights can be written as

$$x = \frac{a_2}{a_1} \leq \frac{q_- + \frac{1}{2}}{2 - q_-}.$$

Since $a_2 > a_1$ and hence $x > 1$, this implies

$$\frac{q_-}{1 - q_-} = x^\alpha \leq x^7 \leq \left(\frac{q_- + \frac{1}{2}}{2 - q_-} \right)^7,$$

and recall that $q_- \geq 1 - \frac{1}{\alpha} > 1/2$.

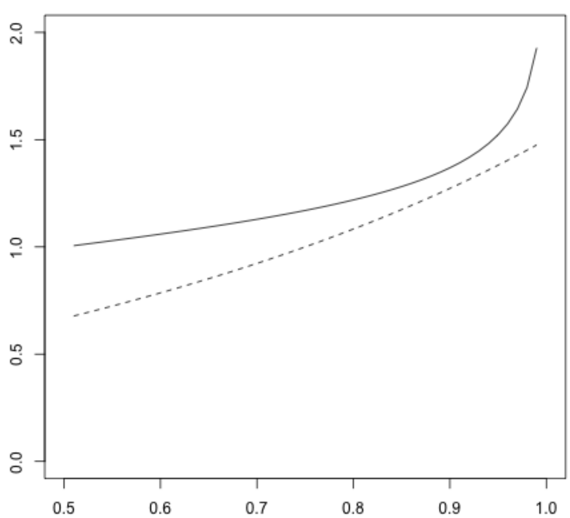


Figure 5: The graphs of functions $y = (\frac{q}{1-q})^{1/7}$ (solid) and $y = (\frac{0.5+q}{2-q})$ (dashed), plotted over $q \in (\frac{1}{2}, 1)$, cut off at $y = 2$.

However, it is not difficult to check that for all $q \in (\frac{1}{2}, 1)$ one has (see Fig. 5)

$$\left(\frac{q}{1-q}\right)^{1/7} > \frac{q+\frac{1}{2}}{2-q}, \tag{27}$$

giving a contradiction. To verify (27) denote the functions on the left and right of (27) by $g_L(q)$ and $g_R(q)$ respectively. Both of these functions are strictly increasing and continuous. As $g_R(1) = \frac{3}{2}$ is finite, we automatically have the inequality (27) for $q \geq q_0 := g_L^{-1}(g_R(1))$. Since $g_R(q_0) < g_R(1) = g_L(q_0)$ we can repeat the above to get that the inequality (27) holds for $q > q_1 := g_L^{-1}(g_R(q_0))$. Iterating this argument, we see that it also holds for $q > F^{(k)}(1)$ for any number of iterations k, where $F(q) = g_L^{-1}(g_R(q))$. Now note that $F^{(5)}(1) < \frac{1}{2}$, hence proving (27) on all the interval $(\frac{1}{2}, 1)$ and thus concluding the proof of the lemma. ■

To handle the $\alpha > 7$ case, we have only to prove the following lemma.

Lemma 5. *Let $\alpha > 7$ and $q \in (0, \frac{1}{2})$. If $\varphi(q) \geq -1$ then $\alpha q < 3 - \sqrt{2}$.*

Proof. The inequality $\varphi(q) \geq -1$ on $(0, \frac{1}{2})$ is equivalent to the inequality $q \leq q_*$, where q_* is the root of the equation $\varphi(q_*) = -1$ in this interval. This equation is quadratic,

$$\alpha q_*^2 - (\alpha - 1)q_* + 1 = 0, \tag{28}$$

and its root in $(0, \frac{1}{2})$ is given by

$$q_*(\alpha) = \frac{\alpha - 1 - \sqrt{\alpha^2 - 6\alpha + 1}}{2\alpha}. \tag{29}$$

Now, it suffices to show that the product $\alpha q_*(\alpha) = \frac{\alpha-1-\sqrt{\alpha^2-6\alpha+1}}{2}$ is decreasing on $\alpha \in (7, \infty)$. Indeed, this will imply the desired

$$\alpha q \leq \alpha q_*(\alpha) \leq 7 \cdot q_*(7) = \frac{6 - \sqrt{8}}{2} = 3 - \sqrt{2} = C.$$

This monotonicity can be easily checked by showing that the first derivative is negative. ■

A.1 Proof of Proposition 10

Recall that (ii) of the Proposition is already proved in [8].

Proof of Proposition 10(i). In the proof, we distinguish between symmetric and asymmetric equilibria. Firstly, [8, proof of Theorem 8] in the special case $r = 1$ excludes *symmetric* fully supported equilibria on $[0, 3]$ when $\alpha < \alpha^*$.

For the asymmetric case, let u be the weight of the interior edge. The equation for the edge incident to a leaf is

$$x = \frac{1}{4} + \frac{1}{4}\frac{x^\alpha}{x^\alpha + u^\alpha}. \tag{30}$$

This can be rearranged to

$$f(x) := 4x^{\alpha+1} - 2x^\alpha + 4u^\alpha x - u^\alpha = 0,$$

and

$$u^\alpha = \frac{2x^\alpha(1-2x)}{4x-1}. \tag{31}$$

We assert that for $\alpha \in [1, \sqrt{8}+3)$ the function $f(x)$ is strictly increasing in $x \in (1/4, 1/2)$. This implies that for such α there is at most one x satisfying (30), so there are no asymmetric equilibria. This will complete the proof of the proposition since $\alpha^* < \sqrt{8}+3$.

To prove the assertion, observe that

$$f'(x) = 4(\alpha+1)x^\alpha - 2\alpha x^{\alpha-1} + 4u^\alpha.$$

Inserting (31) for u^α gives

$$\frac{4x-1}{2x^{\alpha-1}} f'(x) = 2(\alpha+1)(4x-1)x - \alpha(4x-1) + 4x(1-2x).$$

Since $\frac{4x-1}{2x^{\alpha-1}} > 0$, $f'(x) > 0$ if and only if

$$2(\alpha+1)(4x-1)x - \alpha(4x-1) + 4x(1-2x) > 0.$$

Rearranging gives

$$g(x) := \alpha(8x^2 - 6x + 1) + 2x > 0.$$

Now, for fixed α, g has a minimum at $x_0 = \frac{3\alpha-1}{8\alpha}$. Substituting this value of x back into $g(x)$ and simplifying gives $g(x_0) > 0$ if and only if $-\alpha^2+6\alpha-1 > 0$. The roots of this quadratic are $-\sqrt{8}+3 \approx 0.172$ and $\sqrt{8}+3 \approx 5.828$. Hence, if $\alpha \in [1, \sqrt{8}+3)$, then $g(x) > 0$ for all x and hence $f'(x) > 0$. ■

A.2 Proof of Proposition 11

Proposition 11 is a consequence of the following result from [11]:

Theorem 5 ([11])**.** *For $\alpha = 1$ the right hand side of* (20) *is a linearly stable equilibrium for the graph* $[0, 2k]$ *when* $\alpha = 1$*.*

For the reader's benefit we present a brief summary of the proof of this result.

Sketch proof of Theorem 5. One can directly check that the vector on the right hand side of (20) is an equilibrium for $\alpha = 1$. Moreover, for $\alpha = 1$ it is the *unique* fully supported equilibrium, and is linearly stable: the matrix of second derivatives of L at this point is negative definite.

Indeed, the Lyapunov function L (see e.g. (21)) for $\alpha = 1$, is a sum of an affine function $\sum_i x_i$ (that does not affect convexity) and of a linear combination (indexed by vertices v) of non-strictly convex functions $\log\left(\sum_{e_i \sim v} x_i\right)$. For each of these functions, its second derivative in any direction is strictly negative, unless this direction is included in the hyperplane $\sum_{e_i \sim v} x_i = \text{const}$. However, it is easy to check that in our case the intersection of such hyperplanes consists of a single point: for the line graph, we have equations of the type $x_i + x_{i+1} = \text{const}$ as well as $x_1 = \text{const}$ and $x_{2k} = \text{const}$ from the endpoints. Hence, the second derivative of L at the equilibrium point in any direction is strictly negative, and thus the equilibrium for $\alpha = 1$ is linearly stable. ■

Proof of Proposition 11. By Theorem 5 the right hand side of (20) is a fully supported stable equilibrium on $[0, 2k]$ in the case $\alpha = 1$.

Note that we may view the equilibrium equations as equations in both $\{x_e\}_{e \in E}$ and α. Now, a non-degenerate minimum of a function L (or a linearly stable equilibrium of the corresponding anti-gradient flow F) cannot be destroyed by a small perturbation (for the flow F, it follows from the implicit function theorem). Hence, for α sufficiently close to 1, there is also a fully supported equilibrium close to the aforementioned one, that is still linearly stable (due to the continuity of the derivative). ■

Acknowledgments

CH's research is supported by the Centre for Stochastic Geometry and Advanced Bioimaging, funded by grant 8721 from the Villum Foundation. MH's research is supported by Future Fellowship FT160100166 from the Australian Research Council. VK's research is partially supported by the project ANR Gromeov (ANR-19-CE40-0007), as well as by by the Laboratory of Dynamical Systems and Applications NRU HSE, of the Ministry of science and higher education of the RF grant ag. No. 075-15-2019-1931.

References

[1] B. Arthur, Y. Ermoliev, and Y. Kaniovskii. Path dependent processes and the emergence of macro-structure. *Eur. J. Oper. Res.* 30:294–303, (1987).

[2] M. Benaïm, I. Benjamini, J. Chen, and Y. Lima. A generalized Pólya's urn with graph based interactions. *Random Struct. Alg.*, 46(4):614–634, (2015).

[3] M. Benaïm. On strict convergence of stochastic gradients. arXiv:1610.03278 (2018).

[4] B. Davis. Reinforced random walk. *Prob. Theory Related Fields*, 84:203–229, (1990).

[5] Y. Couzinié and C. Hirsch. Weakly reinforced Pólya urns on countable networks. *arxiv:2010.03347* (2020).

[6] C. Hirsch, M. Holmes and V. Kleptsyn. Absence of WARM percolation in the very strong reinforcement regime. *To appear in Ann. Appl. Prob.* (2020).

[7] J. Hofbauer and K. Sigmund. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, (1998).

[8] R.v.d. Hofstad, M. Holmes, A. Kuznetsov, and W. Ruszel. Strongly reinforced Pólya urns with graph-based competition. *Ann. Appl. Probab.*, 26(4):2494–2539 (2016).

[9] C. Hirsch, M. Holmes, and V. Kleptsyn. WARM percolation on a regular tree in the strong reinforcement regime. *preprint*

[10] M. Holmes and V. Kleptsyn. Proof of the WARM whisker conjecture for neuronal connections. *Chaos: an interdisciplinary journal of nonlinear science.* 27. (2017)

[11] M. Holmes and V. Kleptsyn. Infinite WARM graphs II: Critical regime. *in preparation* (2020).

[12] K. Khanin and R. Khanin. A probabilistic model for establishment of neuron polarity. *J. Math. Biol.*, 42:26–40, (2001).

[13] M. Launay and V. Limic. Generalized interacting urn models. http://arxiv.org/abs/1207.5635, (2012).

[14] R. Pemantle. A survey of random processes with reinforcement. *Probability Surveys* 4:1–79, (2007).

[15] G. Pólya. Sur quelques points de la théorie des probabilités *Ann. Inst. Henri Poincaré*, 1(2):117– 161, (1931).

[16] V. B. Tadić. Convergence and convergence rate of stochastic gradient search in the case of multiple and non-isolated extrema. *Stoch. Proc. Appl.* 125:1715–1755, (2015).