The trace-reinforced ants process does not find shortest paths

Daniel Kious*† Cécile Mailler*‡ Bruno Schapira§

January 31, 2022

Abstract

In this paper, we study a probabilistic reinforcement-learning model for ants searching for the shortest path(s) between their nest and a source of food. In this model, the nest and the source of food are two distinguished nodes N and F in a finite graph $\mathcal{G}$. The ants perform a sequence of random walks on this graph, starting from the nest and stopped when first hitting the source of food. At each step of its random walk, the n-th ant chooses to cross a neighbouring edge with probability proportional to the number of preceding ants that crossed that edge at least once. We say that *the ants find the shortest path* if, almost surely as the number of ants grow to infinity, almost all the ants go from the nest to the source of food through one of the shortest paths, without losing time on other edges of the graph.

Our contribution is three-fold: (1) We prove that, if $\mathcal{G}$ is a tree rooted at N whose leaves have been merged into node F, and with one edge between N and F, then the ants indeed find the shortest path. (2) In contrast, we provide three examples of graphs on which the ants do not find the shortest path, suggesting that in this model and in most graphs, ants do not find the shortest path. (3) In all these cases, we show that the sequence of normalised edge-weights converge to a *deterministic* limit, despite a linear-reinforcement mechanism, and we conjecture that this is a general fact which is valid on all finite graphs. To prove these results, we use stochastic approximation methods, and in particular the ODE method. One difficulty comes from the fact that this method relies on understanding the behaviour at large times of the solution of a non-linear, multi-dimensional ODE.

1 Introduction and main results

Context: It is believed that ants are able to find shortest paths between their nest and a source of food with no other means of communication than the pheromones they lay behind them. This phenomenon has been observed empirically in the biology literature (see, e.g., [GADP89, DAGP90, MJT+13]), and reinforcement-learning has been proposed as a model for describing it (see, e.g. [MJT+13]). In the survey of [DS04, Chapter 1], this phenomenon is called *stigmergy*: "*ants stimulate other ants by modifying the environment via pheromone trail updating*".

In this paper, we study a probabilistic reinforcement-learning model for this phenomenon of ants finding shortest path(s) between their nest and a source of food. In this model, which was introduced in our previous paper [KMS], the nest and the source of food are two nodes N and F of a finite graph $\mathcal{G}$, and the ants perform successive random walks starting from N and stopped when first hitting F. The distribution of the n-th ant's walk depends on the trajectory of the previous $n-1$ random walks in a way that models ants leaving pheromones on the edges they cross: each edge has a weight which is equal to 1 at the start and which increases by 1 at time n if and only if the n-th ant has deposited pheromones on

*Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY Bath, UK. Email: d.kious/c.mailler@bath.ac.uk

†DK is grateful to EPSRC for support through the grant EP/V00929X/1.

‡CM is grateful to EPSRC for support through the fellowship EP/R022186/1.

§Aix-Marseille Université, CNRS, Centrale Marseille, I2M, UMR 7373, 13453 Marseille, France. Email: bruno.schapira@univ-amu.fr

that edge. The transition probabilities of the random walk of the $(n+1)$-th ant are proportional to the edge-weights at time n. The reinforcement is thus linear.

In [KMS], the ants deposit pheromones only on their way back: i.e. when they come back to the nest after having hit F. Two cases are studied in [KMS]:

- In the "loop-erased" ant process, or model (LE), ants come back to the nest following the loop-erasure of the time-reversed version of their forward trajectory.
- In the "geodesic" ant process, or model (G), they come back following the shortest path between F and N within the trace of their forward trajectory (ties are broken uniformly at random).

The conjecture is that, under these two versions of the model, when time goes to infinity, almost all ants go from N to F through a shortest path, i.e. *the ants indeed find the shortest path(s) between their nest and the food*. This conjecture is proved in [KMS] for model (LE) in the case when $\mathcal{G}$ is a series-parallel graph, and in the model (G) in the case when $\mathcal{G}$ is a five-edge graph called the lozenge.

Main contribution: In this paper, we look at the same model but assuming that ants deposit pheromones on their way forwards to the food, i.e. the weight of an edge increases by one at time n if and only if the n-th ant has crossed this edge at least once on its way from N to F. We call this model the "trace-reinforced" ant process, or model (T). In the biology literature, all cases of ants depositing pheromones on their way forwards, backwards or both are considered (see [DS04, Chapter 1]).

Maybe surprisingly, this small change to the reinforcement rule leads to a drastically different behaviour: indeed, we prove that, in the trace-reinforced ant process, in general, *the ants do not find the shortest path(s) between their nest and the source of food*, except in some very particular cases. Indeed, in Theorem 1.3, we show that on a family of graphs called "tree-like", in which there is a unique edge between N and F (N and F are at distance 1), the ants do find the shortest path. However, one can find graphs, with N and F at distance one, which are not tree-like and such that the ants do not find shortest paths (see Proposition 1.4, Theorem 1.5 and Proposition 1.6).

The fact that ants do not always find shortest paths when depositing pheromones only on their way forward has been observed empirically on ants (see [DS04, Section 1.1.2]): "*The observation of real ant colonies has confirmed that ants that deposit pheromone only when [going forward] to the nest are unable to find the shortest path between their nest and the food source*". Our analysis proves that the model introduced in [KMS] exhibits the same behaviour.

Our second main finding, which might also be surprising at first glance, is that although we use a *linear* reinforcement mechanism (see the discussion below), in all the graphs we consider, the sequence of normalised edge-weights converges to a *deterministic* limit. In fact, we conjecture that this is a general phenomenon, and that on all graphs with no multiple edges linked to node F, the sequence of the normalised weights converges almost surely to a deterministic limit:

Conjecture 1.1. *If, for all $n \geq 0$, for all $e \in E$, we let $W_e(n)$ denote the weight of edge e at time n, then there exists a deterministic family $(\chi_e)_{e\in E}$ such that $(W_e(n)/n)_{e\in E} \to (\chi_e)_{e\in E}$ almost surely when $n \to +\infty$. Furthermore, if in addition the distance between N and F is at least 2, then $\chi_e > 0$ for all $e \in E$.*

Discussion: Other probabilistic reinforcement models inspired by urns exist in the probability literature. As far as we know these models are all self-reinforced random walks models with super-linear reinforcement; see for example Le Goff and Raimond [LGR18] and Erhard, Franco and Reis [EFR19].

The ant process can be seen as a "path formation" model: the quantity of interest is the subgraph of $\mathcal{G}$ obtained by removing from $\mathcal{G}$ all edges whose normalised weight converges to zero. If this limiting graph is different from $\mathcal{G}$, then we say that "some path(s) has formed".

Another model for path formation is the Pólya urns with graph based interactions of Benaïm, Benjamini, Chen and Lima [BBCL15] and its generalisation to WARMs of [VDHHKR16]. In the latter (and in more recent paper on the same model such as Hirsch, Holmes and Kleptsyn [HHK20]), only super-linear reinforcement is considered because it leads to path formation; in contrast, it is believed that, under linear reinforcement, the limiting graph would be equal to $\mathcal{G}$. In [CH21], Couzinié and Hirsch consider the sub-linear WARM model. They show that, for bounded degree (possibly infinite graphs), if the reinforcement is sufficiently weak, or if $\mathcal{G} = \mathbb{Z}$, then the limiting edge-weights exist and are deterministic. In [BBCL15], and later [CL14] and [Lim16], the cases of sub-linear, super-linear and linear reinforcement are considered. In the linear case, if the original graph $\mathcal{G}$ is regular and not bipartite, then the vector of normalised weights converges almost surely to a *non-deterministic* limit.

Given these examples from the literature, it is quite surprising that in the ant process, with linear reinforcement, the vector of normalised weights always converges to a *deterministic* limit. In fact, this difference of behaviour can be observed when looking at the simplest probabilistic model with linear reinforcement: (generalised) Pólya urns. A d-colour Pólya urn of replacement matrix $R = (R_{i,j})_{1\le i,j\le d}$ is defined as follows: at time zero, there is one ball of each colour in the urn, and at every time step, we pick a ball uniformly at random among the balls in the urn, and if its colour was i, we return it to the urn together with $R_{i,j}$ balls of colour j ($\forall 1 \le j \le d$). If R is the identity, it is known (see Markov [Mar17]) that the vector $\hat{U}(n)$ whose coordinates are the number of balls of each colour divided by n converges almost surely to a Dirichlet(1, ..., 1)-distributed random variable. In contrast, if the matrix R is irreducible, then $\hat{U}(n)$ converges almost surely to a deterministic limit (see Janson [Jan04] and the references therein).

Interestingly, in a not-yet-available paper [HK] (whose results have been announced), Holmes and Klepstyn also exhibit a linear reinforcement model which, on some graphs, converges to a deterministic limit. A result that resonates with Conjecture 1.1.

A similarity between this paper, the Pólya urns with graph-based interaction of [BBCL15] and the WARMs of [VDHHKR16] is the method of proof since we also use stochastic approximation. In particular, we use the ODE method for stochastic approximation, and apply results from Benaïm [Ben99] and Pemantle [Pem07]. However, the analysis of these stochastic approximations in the different examples of graphs we consider is quite different from the one in [BBCL15] and [VDHHKR16]. In fact, as we later explain in more details, each of our examples requires an ad-hoc argument, which suggest that proving our conjecture on the convergence of edge-weights in great generality is a difficult open problem.

That said, we prove a general result ensuring that on *any finite graph*, the sequence of normalised edge weights is a stochastic approximation, in a sense which is recalled in Definition 2.3 below, with a vector field F which is Lipschitz on a suitable convex compact subspace of the Euclidean space (see Proposition 2.7).

1.1 Definition of the model and main results

Let $\mathcal{G} = (V, E)$ be a finite (undirected) graph with vertex set V and edge set E, with two distinct marked nodes called N (for "nest") and F (for "food"). We define a sequence $(\mathbf{W}(n) = (W_e(n))_{e\in E})_{n\ge 0}$ of random weights for the edges of $\mathcal{G}$ recursively as follows:

- At time zero, all weights are equal to 1, i.e. $W_e(0) = 1$ for all $e \in E$.
- Given $\mathbf{W}(n)$, we sample a random walk $(X_i^{(n+1)})_{i\ge 0}$ on $\mathcal{G}$ according to the following distribution:
  - the walk starts at node N, i.e. $X_0^{(n+1)} = N$,
  - it stops when first hitting F, i.e. $\mathbb{P}(X_{i+1}^{(n+1)} = F | X_i^{(n+1)} = F) = 1$ for all $i \ge 0$,
  - for all $i \ge 0$, for all $u \in V \setminus \{F\}$, $v \in V$,

$$\mathbb{P}(X_{i+1}^{(n+1)} = v | X_i^{(n+1)} = u) = \frac{W_{\{u,v\}}(n)}{\sum_{u'\sim u} W_{\{u,u'\}}(n)} \mathbf{1}_{\{u,v\}\in E},$$

    where $u' \sim u$ if there is an edge linking the vertices u' and u.

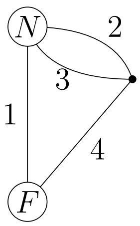

Figure 1: The cone

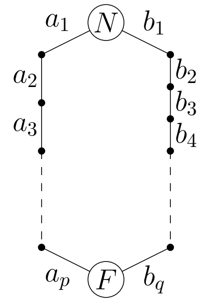

Figure 2: The (p, q)-path graph of Theorem 1.5

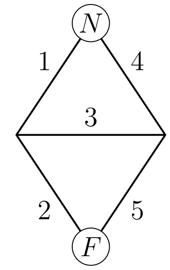

Figure 3: The lozenge

We let $\gamma(n+1)$ be the set of edges that were crossed at least once by the random walk $X^{(n+1)}$; we call this the "trace" of the $(n+1)$-th walker. For all $e \in E$, we set $W_e(n+1) = W_e(n) + \mathbf{1}_{e\in\gamma(n+1)}$.

We call this process the "trace-reinforced" ant process on $\mathcal{G}$.

In our first result, we focus on graphs that are "tree-like" in the following sense:

Definition 1.2. We say that a graph $\mathcal{G} = (V, E)$ with two marked nodes N and F is tree-like if the graph whose vertex set is $V \setminus \{F\}$ and whose edge set is E minus all edges that contain F is a tree (i.e. a graph with no cycle).

Theorem 1.3. *Assume that $\mathcal{G} = (V, E)$ is tree-like and that the edge $a = \{N, F\}$ belongs to E with multiplicity 1. Then, almost surely when $n \to +\infty$,*

$$\frac{W_a(n)}{n} \to 1 \quad \textit{and} \quad \frac{W_e(n)}{n} \to 0, \textit{ for all } e \in E \setminus \{a\}.$$

In other words, following this reinforcement algorithm, the ants eventually find the shortest path between their nest and the source of food, i.e. the proportion of ants that go from N to F by only crossing the edge $\{N, F\}$ is asymptotically equal to one, and the proportion of ants that cross any other edge asymptotically equals zero.

Note that if the edge $\{N, F\}$ appears with multiplicity ℓ in E (i.e. there are ℓ edges from N to F in parallel), then it is easy to deduce from Theorem 1.3 that the normalised weights of all other edges go to zero almost surely, and the weights of the ℓ edges from N to F converge almost surely, as a ℓ-tuple, to a Dirichlet random variable with parameters $(1, \ldots, 1)$.

A natural extension of the set of tree-like graphs is the set of series-parallel graphs, which were considered for instance in [HJ04] and in our previous paper [KMS]. One could then ask whether the previous theorem extends to this class of graphs, that is, if the distance from the source to the food is one, do the weights of all edges not directly connected to both N and F go to zero? Maybe surprisingly, the answer is no. Indeed our next result provides a counter-example, which is depicted by Figure 1 and which we call the *cone* graph.

Proposition 1.4. *Let $\mathcal{G}$ be the graph of Figure 1. If we let $W_i(n)$ be the weight of edge i at time n (using the numbering of edges of Figure 1) and $\mathbf{W}(n) = (W_i(n))_{1\le i\le 4}$, then almost surely when $n \to +\infty$,*

$$\frac{\mathbf{W}(n)}{n} \to (1, 1/3, 1/3, 0).$$

The following result shows that Theorem 1.3 does not extend to tree-like graphs where N and F are at (graph-)distance at least 2 from each other. For two integers $p \geq 1$ and $q \geq 1$, we define the (p, q)-path graph as the graph with two parallel paths between N and F, one of length p and one of length q (see Figure 2).

Theorem 1.5. *Let $\mathcal{G} = (V, E)$ be the (p,q)-path graph. We let $a_1, \dots, a_p$ (resp. $b_1, \dots, b_q$) denote the edges of the path of length p (resp. q), numbered from the closest to the nest to the closest to the food.*

If $\min(p,q) \geq 2$, then, almost surely for all $1 \leq k \leq p$ and $1 \leq \ell \leq q$,

$$\lim_{n\to+\infty} \frac{W_{a_k}(n)}{n} \to \alpha^k \quad \text{and} \quad \lim_{n\to+\infty} \frac{W_{b_\ell}(n)}{n} \to \beta^\ell,$$

where (α, β) is the unique solution in $(0,1)^2$ of

$$\begin{cases} \alpha^p + \beta^q = 1 \\ \alpha^p(1-\alpha) = \beta^q(1-\beta). \end{cases} \tag{1.1}$$

Note that, if $p = q \geq 2$, the solution of (1.1) is explicit and given by $\alpha = \beta = 2^{-1/p}$.

Now to give further support to our conjecture that normalised edge weights always converge to a deterministic limit (when the edges connected to F are simple), we look at the lozenge graph in Figure 3; this example, which was also considered in [KMS], is different from all other cases so far, in the sense that it does not belong to the class of "series-parallel" graphs.

Proposition 1.6. *Let $\mathcal{G}$ be the lozenge graph of Figure 3. If we let $W_i(n)$ denote the weight of edge i at time n (with the edges numbered as in Figure 3), and $\mathbf{W}(n) = (W_i(n))_{1\leq i\leq 5}$, then almost surely as $n \to +\infty$, we have*

$$\frac{\mathbf{W}(n)}{n} \to (w^*, 1/2, 1/2, w^*, 1/2),$$

where w^* is the unique solution of $2x^3 + 4x^2 - 2x - 3/2 = 0$ in $(0,1)$.

The fact that the limiting weights of edges 2 and 5 are equal to $1/2$ should not come as a surprise. Indeed, by symmetry they must be equal (assuming they are deterministic), and since each ant reinforces exactly one of these two edges at each step, their common value has to be $1/2$. Similarly for the edges 1 and 4, except that since each ant can reinforce both of them, one has now $w^* > 1/2$. However, the fact that the limiting weight of edge 3 equals $1/2$ seems to be merely a coincidence.

Notation: Given some filtration $(\mathcal{F}_n)_{n\geq 0}$, and Z some random variable, we will use the notation $\mathbb{E}_n[Z]$ to denote the conditional expectation of Z with respect to $\mathcal{F}_n$.

Acknowledgements: We thank three anonymous referees for their careful reading of the paper and their constructive comments. Preliminary investigations on the problems treated in this paper were carried out by Yassine Hamdi, a student at École Polytechnique at the time, during an undergraduate research internship at the University of Bath, under the supervision of CM (see [Ham20] for Yassine's internship report). The authors are grateful to Yassine for the time he spent on these questions, and to both the École Polytechnique and the University of Bath for making this internship possible.

2 Preliminaries

2.1 Urn processes

We state here a result concerning (generalized) Pólya urn processes. Given a function $G : [0,1] \to [0,1]$, we call *G-urn process*, a process $(X_n)_{n\geq 0}$ with integer values, such that almost surely $X_{n+1} \in \{X_n, X_n + 1\}$, and for all $n \geq 0$,

$$\mathbb{P}(X_{n+1} = X_n + 1 \mid X_0, \dots, X_n) = G(\hat{X}_n),$$

with $\hat{X}_n := \frac{X_n}{n+2}$. In general we will assume that it starts from 1 at time 0, i.e. that $X_0 = 1$, but we shall also consider other initial conditions. We then say that it starts from some value k at time m, if we condition the process on the event $\{X_m = k\}$.

Informally X_n corresponds to the number of (say) red balls after n draws in a Pólya urn with two colours, where at each step, we draw a ball in the urn at random, and replace it into the urn with an additional ball of the same colour. At each draw, the probability to pick a red ball is $G(p)$ if the proportion of red balls in the urn is p.

We will need the following standard result (which follows for instance from Corollary 2.7 and Theorem 2.9 in [Pem07]).

Proposition 2.1. *Let $(X_n)_{n\geq 0}$ be a G-urn process, with G a C^1-function. Then almost surely $(\hat{X}_n)_{n\geq 0}$ converges towards a stable fixed point of G, that is a (possibly random) point $p \in [0,1]$, such that $G(p) = p$ and $G'(p) \leq 1$.*

In particular if there exists $c > 0$, such that $G(x) > x$, for all $x \in (0,c)$ (resp. $G(x) < x$ for all $x \in (1-c,1)$), then almost surely $\liminf_{n\to\infty} \hat{X}_n \geq c$ (resp. $\limsup \hat{X}_n \leq 1-c$).

We shall also use the following corollary.

Corollary 2.2. *Let $(X_n)_{n\geq 0}$ be an integer valued process adapted to some filtration $(\mathcal{F}_n)_{n\geq 0}$, such that almost surely for all $n \geq 0$, $X_{n+1} \in \{X_n, X_n+1\}$, $X_0 = 1$, and for some function $G : [0,1] \to [0,1]$,*

$$\mathbb{P}(X_{n+1} = X_n + 1 \mid \mathcal{F}_n) \geq G(\hat{X}_n), \tag{2.1}$$

with $\hat{X}_n := \frac{X_n}{n+2}$. If there exists $\eta, c > 0$, such that $G(x) > (1+\eta)x$, for all $x \in (0,c)$, then almost surely $\liminf_{n\to\infty} \hat{X}_n \geq c$.

Proof. For $\varepsilon \in (0,\eta)$, consider $G_\varepsilon : [0,1] \to [0,1]$, a C^1 function such that $x < G_\varepsilon(x) \leq (1+\eta)x$, for all $x \in (0, c-\varepsilon)$, $G_\varepsilon(x) \leq (1+\eta)x$, for $x \in (c-\varepsilon, c)$, and $G_\varepsilon \equiv 0$ on $[c,1]$. By assumption on G, one has $G(x) \geq G_\varepsilon(x)$, for all $x \in [0,1]$. It follows that $(X_n)_{n\geq 0}$ stochastically dominates a G_ε-urn process, and applying Proposition 2.1, we deduce that almost surely $\liminf \hat{X}_n \geq c - \varepsilon$. Since this holds for all $\varepsilon \in (0,c)$, the result follows. ☐

In our applications, the process $(X_n)_{n\geq 0}$ will often be one coordinate of a higher-dimensional process $(\mathbf{X}_n)_{n\geq 0}$, and $(\mathcal{F}_n)_{n\geq 0}$ will simply be the natural filtration of the process $(\mathbf{X}_n)_{n\geq 0}$.

2.2 Stochastic approximation and the ODE method

We use the following definition for a stochastic approximation (note that we do not seek for the most general definition here, but it will be sufficient for our purpose).

Definition 2.3. A *stochastic approximation* is a process $(X_n)_{n\geq 0}$, adapted to some filtration $(\mathcal{F}_n)_{n\geq 0}$, with values in a convex compact subset $\mathcal{E} \subseteq \mathbb{R}^d$, for some $d \geq 1$, that satisfies an equation of the type

$$X_{n+1} = X_n + \frac{F(X_n) + \xi_{n+1} + r_n}{n+2}, \qquad \text{for all } n \geq 0, \tag{2.2}$$

where the vector field $F : \mathcal{E} \to \mathbb{R}$ is some Lipschitz function, the *noise* ξ_{n+1} is $\mathcal{F}_{n+1}$-measurable and satisfies $\mathbb{E}_n[\xi_{n+1}] = 0$, for all $n \geq 0$, and the remainder term r_n is $\mathcal{F}_n$-measurable and satisfies almost surely $\|r_n\| \leq C/n$, for some deterministic constant $C > 0$.

Remark. The fact that we assume $\mathcal{E}$ to be a convex compact subset of $\mathbb{R}^d$ enables to easily extend F into a Lipschitz continuous function defined on $\mathbb{R}^d$, simply by composing it with the orthogonal projection on $\mathcal{E}$. We then fall into the setting of Benaïm [Ben99], and we can rely on its results. Thus in the following, we will identify F with its Lipschitz extension on $\mathbb{R}^d$, as defined here.

Remark. The choice of renormalisation factor equal to $\frac{1}{n+2}$ in Equation (2.2) is arbitrary and can be replaced $\frac{1}{n+3}$ (as in Sections 4 and 6). In Sections 4 and 6, we explain why the chosen renormalisation is the most convenient.

The idea underlying the ODE method is that the trajectories of a stochastic approximation *asymptotically follow the solutions* of the differential equation

$$\dot{\boldsymbol{y}} = F(\boldsymbol{y}). \tag{2.3}$$

We recall that if for $x \in \mathbb{R}^d$, we let $(\Phi_t(x))_{t\geq 0}$ be the (unique because F is Lipschitz) solution of (2.3) starting at x, then this defines a *flow*, in the following sense.

Definition 2.4. Let $\mathcal{M}$ be some metric space. A flow (or semi-flow) on $\mathcal{M}$ is an application $\Phi : \mathbb{R}_+ \times \mathcal{M} \to \mathcal{M}$, such that $\Phi_0 = Id$, and $\Phi_{t+s}(x) = \Phi_t \circ \Phi_s(x)$, for all $s, t \geq 0$, and $x \in \mathcal{M}$.

- A subset $\mathcal{A} \subset \mathcal{M}$ is said to be *invariant*, if $\Phi_t(x) \in \mathcal{A}$, for all $x \in \mathcal{A}$ and all $t \geq 0$.
- An *attractor* is a set $\mathcal{A}$ that admits a neighbourhood $\mathcal{U} \subset \mathcal{M}$, such that

$$\cap_{t\geq 0}\overline{\cup_{s>t}\Phi_s(\mathcal{U})} = \mathcal{A}.$$

We will frequently use the following result due to Benaïm [Ben99, Prop. 4.1, Rk. 4.5, Prop. 5.3, Th. 5.7] (see also e.g., [Pem07, Prop. 2.10 and Th. 2.15]).

Theorem 2.5. *Let $(X_n)_{n\geq 1}$ be a stochastic approximation. If there exists a deterministic constant $C > 0$, such that almost surely $\sup_{n\geq 1} \|\xi_n\| \leq C$, then almost surely, the limiting set $L(X) = \cap_{n\geq 0}\overline{\cup_{m\geq n}\{X_m\}}$ is invariant by the flow of the ODE (2.3), connected, and the flow of the ODE restricted to $L(X)$ admits no other attractor than $L(X)$ itself.*

We will also use the following corollary of Theorem 2.5:

Corollary 2.6. *Under the assumptions of Theorem 2.5, if there exist a set $\mathcal{U} \subseteq \mathcal{E}$ and $\boldsymbol{p} \in \mathcal{U}$ such that*

(i) *almost surely, $L(X) \subseteq \mathcal{U}$, and*

(ii) *for all $\boldsymbol{w} \in \mathcal{U}$, the solution of the ODE (2.3) started at $\boldsymbol{w}$ converges to $\boldsymbol{p}$,*

then $L(X) = \{\boldsymbol{p}\}$ almost surely.

Proof. First note that either $L(X) = \{\boldsymbol{p}\}$ or there exists a (possibly random) point $\boldsymbol{x} \in L(X) \setminus \{p\}$. In the first case, the conclusion of Corollary 2.6 holds trivially. In the second case, we show that (i) and (ii) together with Theorem 2.5 imply that $\boldsymbol{p} \in L(X)$. Indeed, by (i), we have that $\boldsymbol{x} \in \mathcal{U}$. Thus, by (ii), the solution $t \mapsto \Phi(t)$ of the ODE started at $\boldsymbol{x}$ converges to $\boldsymbol{p}$ when time goes to infinity. Theorem 2.5 ensures that $L(X)$ is invariant by the flow of the ODE, and thus that $\Phi(t) \in L(X)$ for all $t \geq 0$. Since $L(X)$ is closed (as the intersection of closed sets), this implies that $\lim_{t\to+\infty} \Phi(t) = \boldsymbol{p} \in L(X)$, as claimed. Finally, by Theorem 2.5, the flow restricted to $L(X)$ admits no other attractor than $L(X)$ itself, and by (ii), $\boldsymbol{p}$ is an attractor of the ODE restricted to $L(X) \subseteq \mathcal{U}$; this implies that $L(X) = \{\boldsymbol{p}\}$, as required. □

2.3 The process of edge weights seen as a stochastic approximation

Consider a finite graph $\mathcal{G} = (V, E)$, with two marked vertices N and F. Recall that $\mathbf{W}(n) = (W_e(n))_{e\in E}$ denotes the sequence of the weights of the edges of the graph after n steps of the trace-reinforced ant process, and let $\mathcal{F}_n := \sigma(\mathbf{W}(0), \dots, \mathbf{W}(n))$.

For any edge $e \in E$, and any $n \geq 0$, we let $X_e(n) := \frac{W_e(n)}{n+1}$, and $\boldsymbol{X}(n) = (X_e(n))_{e\in E}$. Next for any $\boldsymbol{w} \in [0, 1]^E$, and any $e \in E$, we let $p_e(\boldsymbol{w})$ be the probability that the edge e belongs to the trace of a random walk on the graph $\mathcal{G}$ endowed with the weights $\boldsymbol{w}$, starting from N and killed at F. Then we define $F : [0, 1]^E \to [0, 1]^E$, by

$$F_e(\boldsymbol{w}) := p_e(\boldsymbol{w}) - w_e, \quad \text{for any } e \in E. \tag{2.4}$$

Given $\boldsymbol{w} \in [0,1]^E$, we set

$$\pi_{\boldsymbol{w}}(x) := \sum_{e\sim x} w_e, \quad \text{for all } x \in V,$$

where $e \sim x$ means that we sum over all edges $e \in E$ that have $x \in V$ as endpoint, and recall that this defines a reversible measure for the random walk on $\mathcal{G}$ endowed with the weights $\boldsymbol{w}$. We also let $\mathfrak{S}(\mathcal{G})$ be the number of self-avoiding paths from N to F in $\mathcal{G}$, which we number in some arbitrary order: $\mathfrak{c}_1, \dots, \mathfrak{c}_{\mathfrak{S}(\mathcal{G})}$. For $i = 1, \dots, \mathfrak{S}(\mathcal{G})$, we define

$$\mathcal{E}_i := \Big\{ \boldsymbol{w} \in [0,1]^E \ : \ \pi_{\boldsymbol{w}}(N) \geq 1, \text{ and } w_e \geq \frac{1}{\mathfrak{S}(\mathcal{G})} \text{ for all } e \in \mathfrak{c}_i \Big\}.$$

Note that each $\mathcal{E}_i$ is a convex compact subset of $[0,1]^E$. Then we further define,

$$\mathcal{E} := \operatorname{conv}\left(\bigcup_{i=1}^{\mathfrak{S}(\mathcal{G})} \mathcal{E}_i \right) = \left\{ \sum_{i=1}^{\mathfrak{S}(\mathcal{G})} \lambda_i \boldsymbol{w}_i \ : \ \begin{array}{l} \sum_i \lambda_i = 1, \text{ and } \lambda_i \geq 0, \text{ for all } i \\ \boldsymbol{w}_i \in \mathcal{E}_i, \text{ for all } i \end{array} \right\}, \tag{2.5}$$

the convex hull of the union of the $\mathcal{E}_i$'s, which is also a convex compact subset of $[0,1]^E$. One has the following general fact.

Proposition 2.7. *The function F is Lipschitz on the space $\mathcal{E}$. Furthermore the process $(\boldsymbol{X}(n))_{n\geq 0}$ is a stochastic approximation on $\mathcal{E}$. More precisely,*

$$\boldsymbol{X}(n+1) = \boldsymbol{X}(n) + \frac{1}{n+2}(F(\boldsymbol{X}(n)) + \boldsymbol{\xi}(n+1)), \tag{2.6}$$

where for any $e \in E$, $\xi_e(n+1) := \mathbf{1}\{W_e(n+1) = W_e(n) + 1\} - p_e(\boldsymbol{X}(n))$.

Remark. We stress that the proof of this result works in a wider setting, including the two variants of the process considered in our previous paper [KMS].

Proof. For the first part, we use a coupling argument. For all $\boldsymbol{w}, \boldsymbol{w}' \in \mathcal{E}$, we define $(X_n)_{n\geq 0}$ as the random walk on $\mathcal{G}$ equipped with edge-weights $\boldsymbol{w}$, and $(X'_n)_{n\geq 0}$ as the random walk on $\mathcal{G}$ equipped with edge-weights $\boldsymbol{w}'$. Both walks start at N and are killed when they first reach F. We couple $(X_n)_{n\geq 0}$ and $(X'_n)_{n\geq 0}$ until the first time when they differ, in a way that maximises the probability that they stay equal after each step. We let τ be the random time when the walks first differ. If τ_F denotes the first time when $(X_n)_{n\geq 0}$ hits F, then

$$\begin{aligned}
\|F(\boldsymbol{w}) - F(\boldsymbol{w}')\|_\infty &= \max_{e\in E} |F_e(\boldsymbol{w}) - F_e(\boldsymbol{w}')| \\
&\leq \max_{e\in E} |p_e(\boldsymbol{w}) - p_e(\boldsymbol{w}')| + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \leq \mathbb{P}(\tau \leq \tau_F) + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \\
&\leq \sum_{k\geq 0} \mathbb{P}(X_k = X'_k, \, X_{k+1} \neq X'_{k+1}, \, k < \tau_F) + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \\
&= \frac{1}{2} \sum_{x\in V} \sum_{k\geq 0} \mathbb{P}(X_k = X'_k = x, \, k < \tau_F) \sum_{e\sim x} \left| \frac{w_e}{\pi_{\boldsymbol{w}}(x)} - \frac{w'_e}{\pi_{\boldsymbol{w}'}(x)} \right| + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty,
\end{aligned} \tag{2.7}$$

where the first inequality is obtained from (2.4) using the triangle inequality, the second inequality is a general fact which holds for any coupling of the two walks, since if an edge is crossed by only one of the two walks, then necessarily $\tau \leq \tau_F$, and the last equality in (2.7) holds because our coupling maximises the probability of the two walks staying equal. From (2.7), we get

$$\begin{aligned}\|F(\boldsymbol{w}) - F(\boldsymbol{w}')\|_\infty &\le \frac{1}{2}\sum_{x\in V}\sum_{k\ge 0}\mathbb{P}(X_k = X'_k = x,\, k<\tau_F)\sum_{e\sim x}\left(\frac{|w_e - w'_e|}{\pi_{\boldsymbol{w}}(x)} + w'_e\frac{|\pi_{\boldsymbol{w}'}(x) - \pi_{\boldsymbol{w}}(x)|}{\pi_{\boldsymbol{w}}(x)\cdot\pi_{\boldsymbol{w}'}(x)}\right)\\ &\quad + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty\\ &\le \frac{1}{2}\sum_{x\in V}\sum_{k\ge 0}\mathbb{P}(X_k = x,\, k<\tau_F)\left(\frac{\sum_{e\sim x}|w_e - w'_e|}{\pi_{\boldsymbol{w}}(x)} + \frac{|\pi_{\boldsymbol{w}'}(x) - \pi_{\boldsymbol{w}}(x)|}{\pi_{\boldsymbol{w}}(x)}\right)\\ &\quad + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty\\ &\le \sum_{x\in V}\frac{G_{\boldsymbol{w}}(N,x)}{\pi_{\boldsymbol{w}}(x)}\cdot\sum_{e\sim x}|w_e - w'_e| + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty,\end{aligned}$$

with $G_{\boldsymbol{w}}(\cdot,\cdot)$ the Green's function on the graph $\mathcal{G}$ endowed with the weights $\boldsymbol{w}$ (i.e. the mean number of visits to the second argument for a random walk starting from the first argument, up to its hitting time of F). Using the reversibility of the measure $\pi_{\boldsymbol{w}}$, we deduce that (see e.g. [LP05, Exercise 2.1(e)]),

$$\frac{G_{\boldsymbol{w}}(N,x)}{\pi_{\boldsymbol{w}}(x)} = \frac{G_{\boldsymbol{w}}(x,N)}{\pi_{\boldsymbol{w}}(N)} \quad (\forall x\in V).$$

Using also that $G_{\boldsymbol{w}}(x,N) \le G_{\boldsymbol{w}}(N,N)$, we get

$$\|F(\boldsymbol{w}) - F(\boldsymbol{w}')\|_\infty \le \frac{G_{\boldsymbol{w}}(N,N)}{\pi_{\boldsymbol{w}}(N)}\sum_{x\in V}\sum_{e\sim x}|w_e - w'_e| + \|\boldsymbol{w} - \boldsymbol{w}'\|_\infty \le \left(1 + \frac{2G_{\boldsymbol{w}}(N,N)}{\pi_{\boldsymbol{w}}(N)}\right)\cdot\|\boldsymbol{w} - \boldsymbol{w}'\|_1,$$

with $\|\boldsymbol{w} - \boldsymbol{w}'\|_1 = \sum_{e\in E}|w_e - w'_e|$. Now by definition, for $\boldsymbol{w}\in\mathcal{E}$, one has $\pi_{\boldsymbol{w}}(N)\ge 1$, and we claim that $G_{\boldsymbol{w}}(N,N)$ is also bounded by a positive constant independent of $\boldsymbol{w}$ (only depending on the graph $\mathcal{G}$). Indeed, by [LP05, Eq. (2.4)], for a random walk starting from N, the number of returns to N before hitting F is a geometric random variable with mean $\pi_{\boldsymbol{w}}(N)/\mathcal{C}_{(\mathcal{G},\boldsymbol{w})}(N,F)$, where $\mathcal{C}_{(\mathcal{G},\boldsymbol{w})}(N,F)$ denotes the effective conductance between N and F in the graph $\mathcal{G}$ endowed with the weights $\boldsymbol{w}$. Moreover, by definition of $\mathcal{E}$, for any $\boldsymbol{w}\in\mathcal{E}$, there exists a self-avoiding path from N to F such that all the edges on this path have a weight larger than $\mathfrak{S}(\mathcal{G})^{-2}$ (we recall that $\mathfrak{S}(\mathcal{G})$ denotes the number of self-avoiding paths between N and F in $\mathcal{G}$). Such a path has an effective conductance larger than $(h_{\max}(\mathcal{G})\cdot\mathfrak{S}(\mathcal{G})^2)^{-1}$, where $h_{\max}(\mathcal{G})$ denotes the maximal length of a self-avoiding path from N to F in $\mathcal{G}$. By Rayleigh's monotonicity principle, we also have $\mathcal{C}_{(\mathcal{G},\boldsymbol{w})}(N,F) \ge (h_{\max}(\mathcal{G})\cdot\mathfrak{S}(\mathcal{G})^2)^{-1}$. Finally, note that $\pi_{\boldsymbol{w}}(N)$ is bounded by the degree of N, say $d_{\mathcal{G}}(N)$. In total, this implies that, for all $\boldsymbol{w},\boldsymbol{w}'\in\mathcal{E}$,

$$\|F(\boldsymbol{w}) - F(\boldsymbol{w}')\|_\infty \le K(\mathcal{G})\cdot\|\boldsymbol{w} - \boldsymbol{w}'\|_1,$$

with $K(\mathcal{G}) := 1 + 2\left(1 + d_{\mathcal{G}}(N)\cdot h_{\max}(\mathcal{G})\cdot\mathfrak{S}(\mathcal{G})^2\right)$, a constant which only depends on the graph $\mathcal{G}$. This concludes the proof of the fact that F is Lipschitz on $\mathcal{E}$.

Since (2.6) is straightforward by definition of the model, it only remains to show that $\boldsymbol{X}(n)$ belongs to $\mathcal{E}$, for all $n\ge 0$. The fact that $\pi_{\boldsymbol{X}(n)}(N)\ge 1$ follows from the fact that, by definition of the model, at each step at least one of the edges incident to N is reinforced. Furthermore, at each step, at least one of the self-avoiding paths from N to F is reinforced, which implies that, at any time $n\ge 0$, at least one of these self-avoiding paths has been reinforced at least $n/\mathfrak{S}(\mathcal{G})$ times. In other words, for all $n\ge 0$, $\boldsymbol{X}(n)$ belongs to at least one of the $\mathcal{E}_i$'s, and thus $\boldsymbol{X}(n)$ belongs to $\mathcal{E}$ as claimed. □

2.4 Case when N and F are at distance one

In this section, we prove the following general fact: if N and F are at distance 1, then the only simple paths from N to F which "survive" asymptotically are those of length one. In other words, asymptotically, almost all of the ants reach F by last crossing one of the edges that connect N and F. However, it is not true in general that the only edges which survive are those from N to F; the cone graph of Proposition 1.4 is a counter-example.

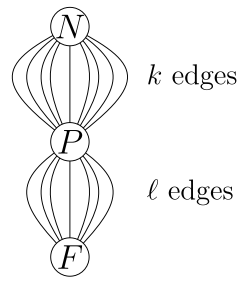

Figure 4: The graph $\mathcal{G}(k,\ell)$.

Proposition 2.8. *Assume that $\mathcal{G}$ is a finite graph with two marked vertices N and F, connected by at least one edge. Let $(\mathbf{W}(n))_{n\geq 0}$ be the process of edge-weights of the trace-reinforced ant-process on $\mathcal{G}$. Then for any edge e connected to F but not to N, one has $W_e(n)/n \to 0$ almost surely.*

Proof. First note that it is enough to prove the result in the case when there is a unique edge $a = \{N, F\}$ (i.e. it has multiplicity one). Indeed, if $\{N, F\}$ has multiplicity $m \geq 2$, then, by definition of the ant process, at most one of these m edges belongs to the trace of each ant. Hence, the process obtained by adding the weights of theses edges into one weight is the ants process on the graph in which the m parallel edges have been merged into one edge with initial weight m.

Assume that $\boldsymbol{w}$ is such that $w_a \notin \{0, 1\}$. Let $\mathcal{G}'$ be the graph obtained by removing edge a from $\mathcal{G}$: i.e. $\mathcal{G}' = (V, E')$ where $E' = E \setminus \{a\}$. We equip the edges of $\mathcal{G}'$ with the weights $(w_e)_{e\in E'}$, and let $\mathcal{C}_{\mathcal{G}'}(\boldsymbol{w})$ denote the conductance between N and F in $\mathcal{G}'$ equipped with these weights. We denote $p_a(\boldsymbol{w})$ the probability to reinforce the edge a when the weights over the graph are given by $\boldsymbol{w}$. With this notation, we have

$$p_a(\boldsymbol{w}) = \frac{w_a}{w_a + \mathcal{C}_{\mathcal{G}'}(\boldsymbol{w})}.$$

We let k denote the number of edges connected to N in $\mathcal{G}'$, and ℓ denote the number of edges connected to F in $\mathcal{G}'$. We define the graph $\mathcal{G}(k,\ell)$ as the graph with vertex set $\{N, F, P\}$ and with edge set $\{N, P\}$ with multiplicity k and $\{P, F\}$ with multiplicity ℓ (see Figure 4). We equip the k edges between N and P in $\mathcal{G}(k,\ell)$ with the same weights as the k edges connected to N in $\mathcal{G}'$, and the ℓ edges between P and F in $\mathcal{G}(k,\ell)$ with the same weights as the ℓ edges connected to F in $\mathcal{G}'$. The graph $\mathcal{G}(k,\ell)$ can be obtained from $\mathcal{G}'$ by merging all vertices different from N and F into one node called P, or equivalently by adding edges between all pair of vertices disjoint of N and F, and assigning an infinite weight to all edges not connected to N or F. By Rayleigh's monotonicity principle (see [LP05, §2.4]), the conductance $\mathcal{C}_{\mathcal{G}(k,\ell)}(\boldsymbol{w})$ of $\mathcal{G}(k,\ell)$ is at least equal to the conductance of $\mathcal{G}'$.

The conductance of $\mathcal{G}(k,\ell)$ with these weights is given by

$$\mathcal{C}_{\mathcal{G}(k,\ell)}(\boldsymbol{w}) = \frac{\displaystyle\sum_{e\in E'\colon N\sim e} w_e \sum_{e\in E'\colon F\sim e} w_e}{\displaystyle\sum_{e\in E'\colon N\sim e} w_e + \sum_{e\in E'\colon F\sim e} w_e} \leq \frac{k(1-w_a)}{k+1-w_a},$$

where $X \sim e$ denotes that the vertex X is an endpoint of e, where we have used that for all $e \in E'$, $w_e \leq 1$, and also that $\sum_{e\in E\colon F\sim e} w_e = 1$, and thus

$$\sum_{e\in E'\colon F\sim e} w_e = 1 - w_a.$$

Therefore,

$$p_a(\boldsymbol{w}) \geq \frac{w_a}{w_a + \frac{k(1-w_a)}{k+1-w_a}}.$$

Thus, $(W_a(n))_{n\geq 0}$ stochastically dominates a G-urn process with

$$G(x) = \frac{x}{x + \frac{k(1-x)}{k+1-x}}.$$

Note that $G(x) = x$ if and only if $x \in \{0,1\}$, and one can compute that $G'(0)= (k+1)/k > 1$. Thus Corollary 2.2 shows that $W_a(n)/n$ converges almost surely to 1, and as a consequence one also has that $W_e(n)/n$ converges almost surely to 0 for all e connected to F different from a. This concludes the proof. □

3 Proof of Theorem 1.3

First note that Proposition 2.8 implies that $W_a(n)/n \to 1$, almost surely. Note also that by definition of the model, it now suffices to show that the normalised weight of all the other edges connected to N go to zero almost surely, as the weight of any edge e in the tree is always smaller than the weight of the unique edge connected to N on the path from e to N.

So let e be some edge connected to N, which is different from a. By assumption, the graph $\mathcal{G}$ is a tree rooted at N and whose leaves have been merged into F. Thus, since the ants are stopped when first hitting F, the first time each ant crosses the edge e has to be from N to the other extremity of e. Moreover, if an ant crosses the edge a it is stopped immediately. This implies that the probability to cross e before reaching F on $\mathcal{G}$ is smaller than on the graph consisting only of the two edges a and e (in parallel between N and F). This gives for all $n \geq 0$, almost surely,

$$\mathbb{P}(W_e(n+1) = W_e(n) + 1 \mid \mathbf{W}(n)) \leq \frac{W_e(n)}{W_e(n) + W_a(n)} = \frac{\hat{W}_e(n)}{\hat{W}_e(n) + \hat{W}_a(n)}, \tag{3.1}$$

where we recall that we write $\hat{W}(n) = W(n)/(n+2)$. Now let $\varepsilon > 0$ be fixed. We know that almost surely, for n large enough, $\hat{W}_a(n) \geq 1 - \varepsilon$. On the other hand on the event when $\hat{W}_a(n) \geq 1 - \varepsilon$, for all n larger than some integer n_0, we know by (3.1) that $(W_e(n))_{n\geq n_0}$ is dominated by a G-urn process with $G(x) = x/(x + 1 - \varepsilon)$. Then Proposition 2.1 shows that almost surely, $\limsup_{n\to\infty} \hat{W}_e(n) \leq \varepsilon$. Since this holds for all $\varepsilon > 0$, this concludes the proof.

4 The cone

We first note that by Proposition 2.7 and the specific features of the cone, the process $\hat{\mathbf{W}}(n) := \mathbf{W}(n)/(n+2)$ is a stochastic approximation on the space

$$\mathcal{E}' = \{\boldsymbol{w} = (w_1, \dots, w_4) \in \mathcal{E} \colon w_1 + w_4 = 1, w_4 \leq w_2 + w_3\},$$

with $\mathcal{E}$ as defined in (2.5). More precisely, for all $n \geq 0$, we have

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \frac{1}{n+3}\big(F(\hat{\mathbf{W}}(n)) + \xi_{n+1}\big),$$

with ξ_{n+1} some martingale difference, and for all $1 \leq i \leq 4$, $F_i(\boldsymbol{w}) = p_i(\boldsymbol{w}) - w_i$, where $p_i(\boldsymbol{w})$ is the probability that edge i belongs to the trace of a random walk on the graph endowed with weights $\boldsymbol{w}$, which starts from N and is killed at F.

Remark. Note that, for the cone graph, it is convenient to define $\hat{\mathbf{W}}(n)$ as $\mathbf{W}(n)/(n+2)$ rather than as $\mathbf{W}(n)/(n+1)$. This is because, by definition of the process, for all $n \geq 0$, $W_1(n) + W_4(n) = n + 2$; indeed, $W_1(0) + W_4(0) = 2$ and, at each step we increase exactly one of $W_1(n)$ and $W_4(n)$ by one.

To calculate p_2 (and thus p_3, by symmetry), we decompose according to the first step of the ant: to reinforce edge 2, it has to either go straight through edge 2 (in which case, no matter what it does later, edge 2 will be reinforced), or go through edge 3. In the latter case, the second step of the ant has to be either through edge 2 (edge 2 gets reinforced no matter what happens next) or back through edge 3, in which case we start again. Hence,

$$p_2(\boldsymbol{w}) = \frac{w_2}{w_1+w_2+w_3} + \frac{w_3}{w_1+w_2+w_3}\left(\frac{w_2}{w_2+w_3+w_4} + \frac{w_3}{w_2+w_3+w_4}\,p_2(\boldsymbol{w})\right).$$

Solving this in terms of $p_2(\boldsymbol{w})$ gives

$$p_2(\boldsymbol{w}) = \frac{w_2(w_2+2w_3+w_4)}{(w_1+w_2+w_3)(w_2+w_3+w_4)-w_3^2}.$$

To calculate $p_1(\boldsymbol{w})$, we use the electrical networks method (see, e.g. [LL10]) to see that the probability, starting from N, to cross edge 1 before edge 4 is w_1 divided by the total conductance between N and F. The latter conductance is given by $w_1 + \frac{(w_2+w_3)w_4}{w_2+w_3+w_4}$, which gives

$$p_1(\boldsymbol{w}) = \frac{w_1}{w_1 + \frac{(w_2+w_3)w_4}{w_2+w_3+w_4}}.$$

Thus, the coordinates of the function F are given by

$$F_1(\boldsymbol{w}) = \frac{w_1}{w_1 + \frac{(w_2+w_3)w_4}{w_2+w_3+w_4}} - w_1 \qquad F_2(\boldsymbol{w}) = \frac{w_2(w_2+2w_3+w_4)}{(w_1+w_2+w_3)(w_2+w_3+w_4)-w_3^2} - w_2$$

$$F_4(\boldsymbol{w}) = -F_1(\boldsymbol{w}) \qquad F_3(\boldsymbol{w}) = \frac{w_3(2w_2+w_3+w_4)}{(w_1+w_2+w_3)(w_2+w_3+w_4)-w_2^2} - w_3.$$

Now our aim is to use the ODE method and for this we proceed in 4 steps:

(1) we first remind that $\lim \hat{W}_1(n) = 1$ almost surely as $n \to +\infty$, thanks to Proposition 2.8,

(2) we prove that $\liminf_{n\to+\infty} \hat{W}_2(n) \wedge \hat{W}_3(n) > 0$ (where $x \wedge y$ denotes the minimum of x and y),

(3) we prove that for any w in $\mathcal{U} := \{\boldsymbol{w} \in \mathcal{E} \colon w_1 = 1, w_2w_3 \neq 0\}$, the solution of $\dot{\boldsymbol{y}} = F(\boldsymbol{y})$ started at $\boldsymbol{w}$, converges to $(1, 1/3, 1/3, 0)$,

(4) we finally apply Corollary 2.6, together with (1), (2) and (3) to conclude.

For **(2)**, note that, if $\boldsymbol{w} \in \mathcal{E}'$, $w_1 \to 1$, $w_2 \to 0$, and $w_4 \to 0$, then

$$(w_1+w_2+w_3)(w_2+w_3+w_4) - w_3^2 \sim w_2+w_3+w_4+w_2w_3+w_3w_4 \sim w_2+w_3+w_4,$$

because $w_2w_3 = o(w_3)$ and $w_3w_4 = o(w_3)$. This implies

$$F_2(\boldsymbol{w}) \sim \frac{w_2(w_2+2w_3+w_4)}{w_2+w_3+w_4} - w_2 = \frac{w_2(w_2+2w_3+w_4) - w_2(w_2+w_3+w_4)}{w_2+w_3+w_4} = \frac{w_2w_3}{w_2+w_3+w_4}.$$

Finally, since, for all $\boldsymbol{w} \in \mathcal{E}'$, $w_4 \le w_2 + w_3$, we get that, as $w_1 \to 1$, $w_4 \to 0$ and $w_2 \to 0$,

$$F_2(\boldsymbol{w}) \ge \frac{w_2w_3(1+o(1))}{2(w_2+w_3)} \ge \frac{w_2 \wedge w_3}{4}(1+o(1)).$$

In other words, there exists $\varepsilon > 0$ such that, for all $\boldsymbol{w} \in \mathcal{E}'$ with $w_2 \le \varepsilon$, and $w_1 \ge 1-\varepsilon$,

$$F_2(\boldsymbol{w}) \ge \frac{w_2 \wedge w_3}{8}. \tag{4.1}$$

By symmetry, for all $\boldsymbol{w} \in \mathcal{E}'$ such that $w_3 \le \varepsilon$, and $w_1 \ge 1-\varepsilon$,

$$F_3(\boldsymbol{w}) \ge \frac{w_2 \wedge w_3}{8}. \tag{4.2}$$

Thus, if we set $H(x) = 9x\mathbf{1}_{x\le\varepsilon}/8$, by Equations (4.1), and (4.2), and since $\hat{W}_1(n) \to 1$ almost surely as $n \to \infty$, we get that, for all sufficiently large n, the random variable $\Phi(n) := W_2(n) \wedge W_3(n)$ satisfies almost surely on the event when $W_2(n) \neq W_3(n)$,

$$\mathbb{P}\big(\Phi(n+1) = \Phi(n) + 1 \mid \mathbf{W}(n)\big) \geq H\big(\hat{\Phi}(n)\big), \tag{4.3}$$

with $\hat{\Phi}(n) := \Phi(n)/n$. This is however not sufficient to apply Corollary 2.2, since when $W_2(n) = W_3(n)$, one has $\Phi(n+1) = \Phi(n) + 1$, only when both edges 2 and 3 are reinforced at the next step, which holds with smaller probability. To overcome this issue, we introduce the following quantity:

$$\Psi(n) := \Phi(n) + \sum_{k=0}^{n-1} \mathbf{1}\{W_2(k) = W_3(k),\, W_2(k+1) \neq W_3(k+1)\}.$$

Note that contrarily to Φ, the function Ψ increases by one unit as soon as at least one of the two edges 2 or 3 is reinforced, whatever their weights are at this time (in particular this holds even when they are equal). Therefore, (4.1) and (4.2) imply that, almost surely for all sufficiently large n,

$$\mathbb{P}\big(\Psi(n+1) = \Psi(n) + 1 \mid \mathbf{W}(n)\big) \geq H\big(\hat{\Phi}(n)\big). \tag{4.4}$$

We now claim that almost surely one has $\Phi(n) \sim \Psi(n)$ as $n \to +\infty$, which by Corollary 2.2 implies $\liminf \Phi(n) > 0$, because $H'(0) = 9/8 > 0$.

In fact we prove a stronger statement: almost surely for all n large enough,

$$\Psi(n) \leq \Phi(n) + \Phi(n)^{3/4}. \tag{4.5}$$

To see this, set

$$Z_n := \max(W_2(n), W_3(n)) - \min(W_2(n), W_3(n)) \quad (\forall n \geq 1).$$

Note that, conditionally on the event that edge 2 or 3 is reinforced but not both, the one with largest weight is more likely to be reinforced than the other. This implies that the process $(Z_k)_{k\geq 0}$ taken at its jump times (i.e. the times $k = 0$ and all $k \geq 1$ such that $Z_k \neq Z_{k-1}$) stochastically dominates a simple random walk $(\mathrm{rSRW}_n)_{n\geq 0}$ on $\mathbb{Z}_+ = \{0, 1, \dots\}$ reflected at 0. We let $L(n)$ denote the number of times $(Z_k)_{k\geq 0}$ returns to zero before time n, i.e.

$$L(n) = \sum_{k=0}^{n} \mathbf{1}\{Z_k = 0\},$$

and $N(n)$ the number of jump times of $(Z_k)_{k\geq 0}$ during its first $L(n)$ excursions out of zero (equivalently the number of times only one of the two edges 2 or 3 is reinforced before the last time before n when the weights of edges 2 and 3 are equal). Then, for all integers n, $L(n)$ is stochastically dominated by the number of returns to the origin of $(\mathrm{rSRW}_k)_{k\geq 0}$ before time $N(n)$. Moreover, by definition,

$$\Psi(n) - \Phi(n) = \sum_{k=0}^{n-1} \mathbf{1}\{W_2(k) = W_3(k),\, W_2(k+1) \neq W_3(k+1)\} \leq 1 + L(n). \tag{4.6}$$

During each excursion out of the origin, the probability to hit level $N^{5/8}$ is equal to $N^{-5/8}$, by a standard Gambler's ruin estimate. Moreover, by Hoeffding's inequality, for any $N \geq 1$, the probability for a simple random walk started at level $N^{5/8}$ to hit 0 before time N is bounded by $2N\exp(-N^{1/4}/2)$. Therefore the probability that $(\mathrm{rSRW}_k)_{k\geq 0}$ returns more than $N^{3/4}/2$ times to the origin before time N is bounded by $(1 - N^{-5/8})^{N^{3/4}/2} + 2N\exp(-N^{1/4}/2) \leq \exp(-N^{1/8}/4)$, for all large N. Using the Borel-Cantelli lemma,

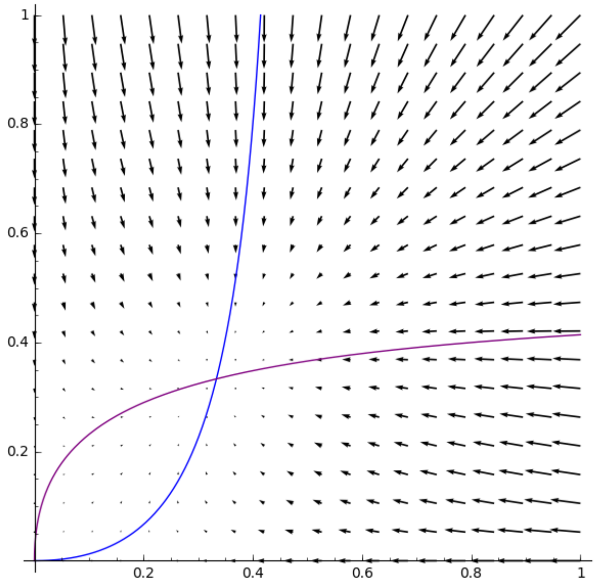

Figure 5: The vector field $F(\boldsymbol{w})$ on $\mathcal{E} \cap \{w_1 = 1\}$ in the cone case.

we deduce that, almost surely for all large N, $(\mathrm{rSRW}_k)_{k \geq 0}$ returns at most $N^{3/4}/2$ times to the origin before time N. This implies that, almost surely for all large n,

$$L(n) \leq \frac{1}{2} N(n)^{3/4}. \tag{4.7}$$

On the other hand, during each excursion of the process $(Z_k)_{k \geq 0}$ out of zero, the process $(\Phi(k))_{k \geq 0}$ increases by at least half the number of jumps made by $(Z_k)_{k \geq 0}$ during this excursion. It follows that $N(n) \leq 2\Phi(n)$, for all $n \geq 0$, and together with (4.6) and (4.7), this concludes the proof of (4.5), and thus of point (**2**).

(3) Let $\boldsymbol{w} \in \mathcal{U}$, and let $\boldsymbol{\Phi}(t) = (\Phi_1(t), \Phi_2(t), \Phi_3(t), \Phi_4(t))$ be the solution of $\dot{\boldsymbol{y}} = F(\boldsymbol{y})$ started at $\boldsymbol{w}$. We need to prove that $\boldsymbol{\Phi}(t) \to (1, 1/3, 1/3, 0)$.

For all $\boldsymbol{w} \in \mathcal{U}$, we have

$$F_2(\boldsymbol{w}) = \frac{w_2(w_2 + 2w_3)}{(1 + w_2 + w_3)(w_2 + w_3) - w_3^2} - w_2 \quad \text{and} \quad F_3(\boldsymbol{w}) = \frac{w_3(2w_2 + w_3)}{(1 + w_2 + w_3)(w_2 + w_3) - w_2^2} - w_3.$$

Note that $F_2(\boldsymbol{w}) = 0$ if and only if $w_2 = 0$ or

$$w_2 + 2w_3 = (1 + w_2 + w_3)(w_2 + w_3) - w_3^2 \quad \Leftrightarrow \quad w_3 = w_2^2 + 2w_2 w_3.$$

Similarly, $F_3(\boldsymbol{w}) = 0$ if and only if $w_3 = 0$ or $w_2 = w_3^2 + 2w_2w_3$. Thus, for all $\boldsymbol{w} \in \mathcal{E} \cap \{w_1 = 1\}$, $F(\boldsymbol{w}) = 0$ if and only if $w_2 = w_3 = 0$ or $w_2 w_3 \neq 0$ and

$$\begin{cases} w_3 = w_2^2 + 2w_2w_3 \\ w_2 = w_3^2 + 2w_2w_3 \end{cases} \quad \Leftrightarrow \quad \begin{cases} w_3 - w_2 = w_2^2 - w_3^2 \\ w_2 = w_3^2 + 2w_2w_3 \end{cases} \quad \Leftrightarrow \quad \begin{cases} w_3 = w_2 \\ 1 = 3w_2 \end{cases}$$

i.e. $w_2 = w_3 = 1/3$. Thus the only zeros of F on $\mathcal{E}' \cap \{w_1 = 1\}$ are $(1, 0, 0, 0)$ and $(1, 1/3, 1/3, 0)$. Similar calculations show that $F_2(\boldsymbol{w}) > 0$ if and only if $w_2 < w_3^2/(1 - 2w_3)$, and $F_3(\boldsymbol{w}) > 0$ if and only if $w_3 < w_2^2/(1 - 2w_2)$. In Figure 5, we plot the vector field $(F_2(1, w_2, w_3, 0), F_3(1, w_2, w_3, 0))$ with w_2 on the horizontal axis and w_3 on the vertical axis. The blue curve is where $F_2 = 0$, the purple curve where

$F_3 = 0$. Note that $[0,1]^2 \setminus (\{w_3 = w_2^2/(1-2w_2)\} \cup \{w_2 = w_3^2/(1-2w_3)\})$ has four connected components: on the bottom-left one, $F_2, F_3 > 0$, on the top-right one, $F_2, F_3 < 0$, on the bottom right one, $F_2 < 0$ while $F_3 > 0$, and, finally, on the top-left one, $F_2 > 0$ while $F_3 < 0$.

Thus, any solution of the ODE started in $\mathcal{U}$ converges to $(1, 1/3, 1/3, 0)$, as claimed.

(4) Finally, we prove that $\hat{\mathbf{W}}(n) \to (1, 1/3, 1/3, 0)$ almost surely as $n \to +\infty$. For this we apply Corollary 2.6: in Steps (1) and (2), we have shown that $\lim_{n\to+\infty} \hat{W}_1(n) = 1$ and $\liminf_{n\to+\infty} \hat{W}_2(n) \wedge \hat{W}_3(n) > 0$. This implies that almost surely the limiting set $L(\mathbf{W})$ of the process $(\hat{\mathbf{W}}(n))_{n\geq 0}$ is contained in $\mathcal{U}$. Then **(3)** and Corollary 2.6 imply that $L(\mathbf{W}) = \{(1, 1/3, 1/3, 0)\}$, as wanted.

5 Two parallel paths: proof of Theorem 1.5

Recall that, in Theorem 1.5, the graph $\mathcal{G} = (V, E)$ is the $(p+q)$-path graph of Figure 2. In this section, we assume that $\min(p, q) \geq 2$. We let $a_1, \dots, a_p$ denote the p edges on one of the paths from N to F (ordered from N to F, i.e. a_1 links to N while a_p links to F), and $b_1, \dots, b_q$ denote the edges on the other path from N to F (also ordered from N to F).

Finally, for all $e \in E = \{a_1, \dots, a_p, b_1, \dots, b_q\}$, we let $W_e(n)$ denote the weight of edge e at time n and set $\mathbf{W}(n) = (W_{a_1}(n), \dots, W_{a_p}(n), W_{b_1}(n), \dots, W_{b_q}(n))$, for all $n \geq 0$.

The proof of Theorem 1.5 uses the ODE method, as for the cone graph in the previous section. We roughly follow the same steps here, but there are some important differences. We first show in Subsection 5.1 that almost surely the limiting set $L(\mathbf{W})$ of the sequence of normalised weights $\hat{\mathbf{W}}(n) := \mathbf{W}(n)/(n+1)$, is contained in $(0,1]^{p+q}$. Then, in Subsection 5.2, we give an explicit expression for the vector field F appearing in the stochastic approximation satisfied by $(\hat{\mathbf{W}}(n))_{n\geq 0}$. In this section, we also define a sequence of compact sets $(K_n)_{n\geq 0}$ in $[0,1]^{p+q}$, which we prove to be decreasing. Then in Subsection 5.3, we show that $L(\mathbf{W}) \subseteq K_n$, for all $n \geq 0$, and finally in Subsection 5.4, we prove that the intersection of the K_n's is reduced to a single point, which is precisely the limiting point arising in the statement of Theorem 1.5.

5.1 Proof that none of the edges has a limiting weight equal to zero

We prove here the following result.

Proposition 5.1. *Let p and q be integers such that $\min(p, q) \geq 2$. If $\mathcal{G} = (V, E)$ is the (p, q)-path graph, then there exists a constant $c_{p,q} > 0$, such that almost surely for all $e \in E$,*

$$\liminf_{n\to+\infty} \frac{W_e(n)}{n} > c_{p,q}.$$

This proposition is proved by induction on k for $a_k \in \{a_1, \dots, a_p\}$ and by induction on ℓ for $b_\ell \in \{b_1, \dots, b_q\}$. We prove the base case separately in the following lemma.

Lemma 5.2. *In the $(p+q)$-path graph, if $q \geq 2$, then there exists $c > 0$, such that almost surely, $\liminf_{n\to+\infty} W_{a_1}(n)/n \geq c$, and by symmetry, if $p \geq 2$, then almost surely, $\liminf_{n\to+\infty} W_{b_1}(n)/n \geq c$.*

Proof. For all $n \geq 0$, we have

$$\mathbb{P}(W_{a_1}(n+1) = W_{a_1}(n) + 1 \mid \mathbf{W}(n)) = \mathbb{P}(a_1 \in \gamma(n+1) \mid \mathbf{W}(n)) = \frac{W_{a_1}(n)}{W_{a_1}(n) + \frac{1}{\sum_{i=1}^{q} W_{b_i}(n)^{-1}}}. \qquad (5.1)$$

Note that, by definition of the model, for all $n \geq 0$, we have either $a_p \in \gamma(n+1)$ or $b_q \in \gamma(n+1)$ but not both. This implies that, for all $n \geq 0$,

$$W_{a_p}(n) + W_{b_q}(n) = n + 2.$$

By definition of the model again, we have that $a_p \in \gamma(n+1) \Rightarrow a_{p-1} \in \gamma(n+1) \Rightarrow \cdots \Rightarrow a_1 \in \gamma(n+1)$, and $b_q \in \gamma(n+1) \Rightarrow b_{q-1} \in \gamma(n+1) \Rightarrow \cdots \Rightarrow b_1 \in \gamma(n+1)$, which imply that, for all $n \geq 0$,

$$W_{a_p}(n) \leq W_{a_{p-1}}(n) \leq \cdots \leq W_{a_1}(n) \leq n+1 \quad \text{and} \quad W_{b_q}(n) \leq W_{b_{q-1}}(n) \leq \cdots \leq W_{b_1}(n) \leq n+1.$$

Thus, (5.1) implies that

$$\mathbb{P}(W_{a_1}(n+1) = W_{a_1}(n) + 1 \mid \mathbf{W}(n)) \geq \frac{W_{a_1}(n)}{W_{a_1}(n) + \frac{1}{q(n+1)^{-1}}} = \frac{Z(n)}{Z(n) + 1/q},$$

where we have set $Z(n) := W_{a_1}(n)/(n+1)$, for all $n \geq 0$. This implies that $(W_{a_1}(n))_{n\geq 0}$ stochastically dominates a G-urn process, where, for all $z \in [0,1]$,

$$G(z) = \frac{z}{z + 1/q}.$$

Note that $G(z) = z$ if and only if $z = 0$ or $z = 1 - 1/q$ (which is positive because $q \geq 2$, by assumption); furthermore, one can check that G is C^1 and $G'(0) = q > 1$, implying that the normalized G-urn process converges almost surely to $c := 1 - 1/q > 0$, when $n \to +\infty$, by Proposition 2.1. This concludes the proof. ☐

Proof of Proposition 5.1. We prove the result by induction on $k \in \{1, \ldots, p\}$. The case $k = 1$ follows from Lemma 5.2. Assume the result holds true for some $k < p$, i.e. there exists $c \in (0, 1/2)$ such that almost surely $\liminf W_{a_k}(n)/n > c$. Let us define the events

$$E_m := \{W_{a_k}(n) \geq cn, \forall n \geq m\} \qquad (m \geq 1).$$

Then, $\bigcup_m E_m$ holds almost surely.

Let τ_n be the time when edge a_k is reinforced for the n-th time (so by definition $W_{a_k}(\tau_n) = n+1$). Note that by definition $\tau_n \geq n$ for all $n \geq 1$. Thus, for all $n \geq m$, on E_m, using that $W_{a_{k-1}}(\tau_n) \leq \tau_n + 1$ and $\tau_n \leq (n+1)/c$, one has

$$\frac{W_{a_k}(\tau_n)}{W_{a_k}(\tau_n) + W_{a_{k-1}}(\tau_n)} \geq \frac{c(n+1)}{c(n+1) + n + 1 + c} \geq c/2.$$

It follows that, almost surely on the event E_m, if the weight of a_{k+1} is zero, then, for all $n \geq m$, the τ_n-th ant makes at least a geometric number of crossings of edge a_k, with success probability $\nu := 1 - c/2$, before jumping across edge a_{k-1}. Consequently, on E_m, and for $n \geq m$, the probability to reinforce edge a_{k+1} at time τ_n is at least $1 - (1 - \rho(n))^{X_n}$, where $(X_n)_{n\geq 0}$ is a sequence of i.i.d. geometric random variables with parameter ν, and

$$\rho(n) = \frac{W_{a_{k+1}}(\tau_n)}{W_{a_{k+1}}(\tau_n) + n + 1}.$$

Thus, letting $u_{k+1}(n) = W_{a_{k+1}}(\tau_n)/(n+1)$, we find

$$\mathbb{E}[W_{a_{k+1}}(\tau_{n+1}) - W_{a_{k+1}}(\tau_n) \mid \mathcal{F}_{\tau_n}] \geq \mathbb{E}[1 - (1-\rho(n))^{X_n} \mid W_{a_{k+1}}(\tau_n)] = 1 - \frac{\nu(1-\rho(n))}{1 - (1-\nu)(1-\rho(n))}$$

$$= \frac{\rho(n)}{1 - (1-\nu)(1-\rho(n))} = \frac{u_{k+1}(n)}{u_{k+1}(n) + \nu},$$

with $(\mathcal{F}_i)_{i\geq 0}$ the natural filtration of the process, and where we used that $\mathbb{E}x^{X_n} = \nu x/(1 - (1-\nu)x)$ for all $x \in (0,1)$. It follows that on the event E_m, the process $(W_{a_{k+1}}(\tau_n))_{n\geq m}$ stochastically dominates a G-urn process (starting from $W_{a_{k+1}}(\tau_m)$), with $G(x) := \frac{x}{x+\nu}$. Since $G(x) > x$ for all $x \in (0, c/2)$, it follows from Lemma 2.1 that almost surely $\liminf u_{k+1}(n) \geq c/2$. We deduce the induction step, using that by hypothesis, $\limsup \tau_n/n \leq 1/c$. ☐

5.2 The stochastic algorithm and a sequence of decreasing compact subspaces.

Recall that we set $\hat{\mathbf{W}}(n) := \mathbf{W}(n)/(n+1)$, for all $n \geq 0$. Note also that, by definition of the model, for all $n \geq 0$,

$$\hat{\mathbf{W}}(n) \in \mathcal{E}' := \{\boldsymbol{w} \in \mathcal{E}\colon w_{a_p} = 1 - w_{b_q}, w_{b_q} \leq w_{b_{q-1}} \leq \cdots \leq w_{b_1}, w_{a_p} \leq w_{a_{p-1}} \leq \cdots \leq w_{a_1}\},$$

with $\mathcal{E}$ as defined in (2.5). Moreover, for all $n \geq 0$, we have

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \frac{1}{n+2}\big(F(\hat{\mathbf{W}}(n)) + \xi_{n+1}\big),$$

with ξ_{n+1} some martingale difference and with F as defined in (2.4). More specifically the coordinates of F can be computed explicitly here, and are given by, for all $1 \leq k \leq p$ and $1 \leq \ell \leq q$, for all $\boldsymbol{w} \in \mathcal{E}'$,

$$F_{a_k}(\boldsymbol{w}) = \frac{S_k^a(\boldsymbol{w})}{S_k^a(\boldsymbol{w}) + S_q^b(\boldsymbol{w})} - w_{a_k} \quad \text{and} \quad F_{b_\ell}(\boldsymbol{w}) = \frac{S_\ell^b(\boldsymbol{w})}{S_\ell^b(\boldsymbol{w}) + S_p^a(\boldsymbol{w})} - w_{b_k}, \tag{5.2}$$

where we have defined, for any $\boldsymbol{w} \in [0,1]^E$, $m \in \{a, b\}$, and s an integer such that $1 \leq s \leq p$ if $m = a$, and $1 \leq s \leq q$ if $m = b$,

$$S_s^m(\boldsymbol{w}) = \frac{1}{\sum_{i=1}^{s} \frac{1}{w_{m_i}}}. \tag{5.3}$$

Note that, for all $1 \leq k \leq p$, $S_k^a(\boldsymbol{w}) = 0$ if and only if $w_{a_i} = 0$ for some $1 \leq i \leq k$, and for all $1 \leq \ell \leq q$, $S_\ell^b(\boldsymbol{w}) = 0$ if and only if $w_{b_i} = 0$ for some $1 \leq i \leq \ell$.

To prove Theorem 1.5, we use the ODE method and thus start by studying the solutions of the equation $\dot{\boldsymbol{y}} = F(\boldsymbol{y})$. To do so, we define a sequence $(K_n)_{n\geq 0}$ of decreasing compact subsets of $\mathcal{E}'$ such that (A) for all $n \geq 0$, $L(\mathbf{W}) \subseteq K_n$, and (B) the intersection of all these compacts is $\{\boldsymbol{w}^*\}$, where $w^*_{a_k} = \alpha^k$ and $w^*_{b_\ell} = \beta^\ell$, with (α, β) as in Theorem 1.5. We prove (A) in Section (5.3), and (B) in Section 5.4. In the rest of this section, we define the sequence $(K_n)_{n\geq 0}$ and show that it is decreasing, i.e. that $K_{n+1} \subset K_n$ for all $n \geq 0$. For this we need some additional notation. For all $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{E}'$, we let

$$H_{a_k}(\boldsymbol{u}, \boldsymbol{v}) = \frac{S_k^a(\boldsymbol{u})}{S_k^a(\boldsymbol{u}) + S_q^b(\boldsymbol{v})} \quad \text{and} \quad H_{b_\ell}(\boldsymbol{u}, \boldsymbol{v}) = \frac{S_\ell^b(\boldsymbol{u})}{S_\ell^b(\boldsymbol{u}) + S_p^a(\boldsymbol{v})},$$

for all $1 \leq k \leq p$ and $1 \leq \ell \leq q$. Recall further that by Proposition 5.1, there exists a constant $c_{p,q} > 0$, such that almost surely

$$L(\mathbf{W}) \subseteq \{\boldsymbol{w}\colon w_e \geq c_{p,q}, \text{ for all } e \in E\}. \tag{5.4}$$

We also define $\boldsymbol{w}^*$ as the limiting vector appearing in the statement of Theorem 1.5. More precisely, we have

$$w^*_{a_k} = \alpha^k, \quad \text{and} \quad w^*_{b_\ell} = \beta^\ell, \tag{5.5}$$

for all $1 \leq k \leq p$, $1 \leq \ell \leq q$, with (α, β) the unique solution of the system (1.1) in $(0,1)^2$ (existence and uniqueness of the solution for this system of equations will be proved later, at the end Subsection 5.4). We then define $\boldsymbol{u}^{(0)}$ and $\boldsymbol{v}^{(0)}$ by

$$u^{(0)}_{a_k} = u_0^k, \quad u^{(0)}_{b_\ell} = u_0^\ell, \quad \text{and} \quad v^{(0)}_{a_k} = v^{(0)}_{b_\ell} = 1,$$

for all $1 \leq k \leq p$, $1 \leq \ell \leq q$, with u_0 chosen arbitrarily so that

$$0 < u_0 < \min(1 - {}^1\!/_q, \alpha, \beta, c_{p,q}).$$

The fact that we choose $u_0 \leq \min(\alpha, \beta)$ entails in particular

$$0 < u^{(0)}_e \leq w^*_e, \quad \text{for all } e \in E.$$

Next we define inductively two sequences $(\boldsymbol{u}^{(n)})_{n\geq 0}$ and $(\boldsymbol{v}^{(n)}))_{n\geq 0}$, by

$$\boldsymbol{u}^{(n+1)} = H(\boldsymbol{u}^{(n)}, \boldsymbol{v}^{(n)}), \quad \text{and} \quad \boldsymbol{v}^{(n+1)} = H(\boldsymbol{v}^{(n)}, \boldsymbol{u}^{(n)}).$$

Finally, we define the sequence $(K_n)_{n\geq 0}$ by

$$K_n := \{\boldsymbol{w} \in \mathcal{E}' \colon u_e^{(n)} \leq w_e \leq v_e^{(n)}, \text{ for all } e \in E\} \quad (\forall n \geq 0).$$

We prove now that this sequence is decreasing, that is, $K_{n+1} \subset K_n$, for all $n \geq 0$. More precisely, we prove that, for all $n \geq 0$,

$$\boldsymbol{u}^{(n)} < H(\boldsymbol{u}^{(n)}, \boldsymbol{v}^{(n)}) = \boldsymbol{u}^{(n+1)} \tag{5.6}$$

$$\boldsymbol{v}^{(n)} > H(\boldsymbol{v}^{(n)}, \boldsymbol{u}^{(n)}) = \boldsymbol{v}^{(n+1)}. \tag{5.7}$$

We reason by induction on n: first note that, for all $1 \leq k \leq p$ and $1 \leq \ell \leq q$,

$$S_k^a(\boldsymbol{u}^{(0)}) = u_0^k \frac{1-u_0}{1-u_0^k} \quad \text{and} \quad S_\ell^b(\boldsymbol{v}^{(0)}) = \frac{1}{\ell},$$

which implies, using that $1 - u_0 > 1/q$,

$$u_{a_k}^{(1)} = H_{a_k}(\boldsymbol{u}^{(0)}, \boldsymbol{v}^{(0)}) = \frac{u_0^k(1-u_0)}{u_0^k(1-u_0) + (1-u_0^k)/q} > \frac{u_0^k/q}{u_0^k/q + (1-u_0^k)/q} = u_0^k = u_{a_k}^{(0)},$$

and

$$v_{a_k}^{(1)} = H_{a_k}(\boldsymbol{v}^{(0)}, \boldsymbol{u}^{(0)}) < 1 = v_{a_k}^{(0)}.$$

Similarly, for all $1 \leq \ell \leq q$,

$$u_{b_\ell}^{(1)} > u_{b_\ell}^{(0)} \quad \text{and} \quad v_{b_\ell}^{(1)} < v_{b_\ell}^{(0)}.$$

We now proceed to the induction step and assume that, for some $n \geq 1$ $u_e^{(n)} > u_e^{(n-1)}$ and $v_e^{(n)} < v_e^{(n-1)}$ for all $e \in E$. We then simply observe that, for all $e \in E$,

$$u_e^{(n+1)} = H_e(\boldsymbol{u}^{(n)}, \boldsymbol{v}^{(n)}) > H_e(\boldsymbol{u}^{(n-1)}, \boldsymbol{v}^{(n-1)}) = u_e^{(n)},$$

and

$$v_e^{(n+1)} = H_e(\boldsymbol{v}^{(n)}, \boldsymbol{u}^{(n)}) < H_e(\boldsymbol{v}^{(n-1)}, \boldsymbol{u}^{(n-1)}) = v_e^{(n)},$$

which proves the induction step, and thus concludes the proofs of (5.6) and (5.7). In other words we just have proved that $(K_n)_{n\geq 0}$ is indeed a sequence of decreasing sets.

5.3 Proof that $L(\mathbf{W}) \subseteq K_n$, for all $n \geq 0$.

We use an induction argument. Note first that by (5.4) and the definition of K_0, one has almost surely $L(\mathbf{W}) \subseteq K_0$, using also the hypothesis $u_0 \leq c_{p,q}$. We now prove the induction step, i.e. that almost surely, if $L(\mathbf{W}) \subseteq K_n$, for some $n \geq 0$, then also $L(\mathbf{W}) \subseteq K_{n+1}$.

To do so, we first look at $F_{a_k}(\boldsymbol{w})$ for $1 \leq k \leq p$ and $\boldsymbol{w}$ such that $u_e \leq w_e \leq v_e$ for all $e \in E$: we have

$$F_{a_k}(\boldsymbol{w}) = \frac{S_k^a(\boldsymbol{w})}{S_k^a(\boldsymbol{w}) + S_q^b(\boldsymbol{w})} - w_{a_k} \geq \frac{S_k^a(\boldsymbol{u})}{S_k^a(\boldsymbol{u}) + S_q^b(\boldsymbol{v})} - w_{a_k} = H_{a_k}(\boldsymbol{u}, \boldsymbol{v}) - w_{a_k}.$$

Thus, if $u_e \leq w_e \leq v_e$ for all $e \in E$ and

$$w_{a_k} < H_{a_k}(\boldsymbol{u}, \boldsymbol{v}),$$

then $F_{a_k}(\boldsymbol{w}) > 0$. Also, for all $1 \leq k \leq p$, if $u_e \leq w_e \leq v_e$ for all $e \in E$ and

$$w_{a_k} > H_{a_k}(\boldsymbol{v}, \boldsymbol{u}),$$

then $F_{a_k}(\boldsymbol{w}) < 0$. The same argument leads to

$$w_{b_\ell} < H_{b_\ell}(\boldsymbol{u},\boldsymbol{v}) \implies F_{b_\ell}(\boldsymbol{w}) > 0;$$
$$w_{b_\ell} > H_{b_\ell}(\boldsymbol{v},\boldsymbol{u}) \implies F_{b_\ell}(\boldsymbol{w}) < 0,$$

for all $1 \le \ell \le q$, and if $u_e \le w_e \le v_e$ for all $e \in E$. These facts imply that, for all $n \ge 0$, and for any $\boldsymbol{w} \in K_n$, the flow of the ODE $\dot{\boldsymbol{y}} = F(\boldsymbol{y})$ started at $\boldsymbol{w}$ converges to K_{n+1}. Therefore if we already know that $L(\mathbf{W}) \subseteq K_n$, then it means that $L(\mathbf{W}) \cap K_{n+1}$ is an attractor of the flow restricted to $L(\mathbf{W})$. Thus by Proposition 2.5, we deduce that $L(\mathbf{W}) \subseteq K_{n+1}$.

Altogether this proves that $L(\mathbf{W})$ is almost surely included in the intersection of all the K_n's.

5.4 Identification of the intersection of the K_n's.

We show here that the intersection of the K_n's is reduced to the single point $\boldsymbol{w}^*$, which appears as the limiting vector in the statement of Theorem 1.5 (see also (5.5) above).

Since the sequences $(u_e^{(n)})$ and $(v_e^{(n)})$ are all monotonic and bounded, they all converge. We let $\boldsymbol{u}^* = \lim_{n\to+\infty} \boldsymbol{u}^{(n)}$ and $\boldsymbol{v}^* = \lim_{n\to+\infty} \boldsymbol{u}^{(n)}$. Because H is continuous on $\mathcal{E}'$, we have

$$\boldsymbol{u}^* = H(\boldsymbol{u}^*,\boldsymbol{v}^*), \quad \text{and} \quad \boldsymbol{v}^* = H(\boldsymbol{v}^*,\boldsymbol{u}^*).$$

The equation $\boldsymbol{u}^* = H(\boldsymbol{u}^*,\boldsymbol{v}^*)$ can be written as, for all $1 \le k \le p$, $1 \le \ell \le q$,

$$u^*_{a_k} = \frac{S^a_k(\boldsymbol{u}^*)}{S^a_k(\boldsymbol{u}^*) + S^b_q(\boldsymbol{v}^*)} \quad \text{and} \quad u^*_{b_\ell} = \frac{S^b_\ell(\boldsymbol{u}^*)}{S^b_\ell(\boldsymbol{u}^*) + S^a_p(\boldsymbol{v}^*)}.$$

Using that $S^a_1(\boldsymbol{u}^*) = u^*_{a_1}$ and $S^b_1(\boldsymbol{u}^*) = u^*_{b_1}$, this implies that

$$u^*_{a_1} = 1 - S^b_q(\boldsymbol{v}^*) =: \alpha \quad \text{and} \quad u^*_{b_1} = 1 - S^a_p(\boldsymbol{v}^*) =: \beta, \tag{5.8}$$

and, for all $2 \le k \le p$, $2 \le \ell \le q$,

$$u^*_{a_k} = \frac{S^a_k(\boldsymbol{u}^*)}{S^a_k(\boldsymbol{u}^*) + 1 - u^*_{a_1}} \quad \text{and} \quad u^*_{b_\ell} = \frac{S^b_\ell(\boldsymbol{u}^*)}{S^b_\ell(\boldsymbol{u}^*) + 1 - u^*_{b_1}}.$$

We first show by induction that this implies $u^*_{a_k} = \alpha^k$ and $u^*_{b_\ell} = \beta^\ell$ for all $1 \le k \le p$ and $1 \le \ell \le q$. Indeed, if for some $1 < k \le p$, and all $1 \le i < k$, $u^*_{a_i} = \alpha^i$, then

$$u^*_{a_k} = \frac{\left(\frac{1}{u^*_{a_k}} + \frac{1-\alpha^{k-1}}{\alpha^{k-1}(1-\alpha)}\right)^{-1}}{\left(\frac{1}{u^*_{a_k}} + \frac{1-\alpha^{k-1}}{\alpha^{k-1}(1-\alpha)}\right)^{-1} + 1 - \alpha} = \frac{1}{1 + (1-\alpha)\left(\frac{1}{u^*_{a_k}} + \frac{1-\alpha^{k-1}}{\alpha^{k-1}(1-\alpha)}\right)},$$

and a straightforward calculation yields that $u^*_{a_k} = \alpha^k$, as claimed. The proof of $u^*_{b_\ell} = \beta^\ell$ for all $1 \le \ell \le q$ is similar.

Since $(\boldsymbol{u}^*,\boldsymbol{v}^*)$ also satisfies the symmetric equation $\boldsymbol{v}^* = H(\boldsymbol{v}^*,\boldsymbol{u}^*)$, we get that

$$v^*_{a_1} = 1 - S^b_q(\boldsymbol{u}^*) =: \bar{\alpha} \quad \text{and} \quad v^*_{b_1} = 1 - S^a_p(\boldsymbol{u}^*) =: \bar{\beta}, \tag{5.9}$$

and $v^*_{a_k} = \bar{\alpha}^k$, $v^*_{b_\ell} = \bar{\beta}^\ell$ for all $1 \le k \le p$ and $1 \le \ell \le q$. Using this into (5.8), and using the definition of S_p and S_q (see (5.3)), we get

$$\alpha = 1 - \frac{1}{\sum_{i=1}^q \bar{\beta}^{-i}} \quad \text{and} \quad \beta = 1 - \frac{1}{\sum_{i=1}^p \bar{\alpha}^{-i}}.$$

Similarly, using the fact that $u^*_{a_k} = \alpha^k$ and $u^*_{b_\ell} = \beta^\ell$ for all $1 \le k \le p$ and $1 \le \ell \le q$, together with Equation (5.9), we get

$$\bar{\alpha} = 1 - \frac{1}{\sum_{i=1}^{q} \beta^{-i}} \quad \text{and} \quad \bar{\beta} = 1 - \frac{1}{\sum_{i=1}^{p} \alpha^{-i}}.$$

Note that

$$\alpha = 1 - \frac{1}{\sum_{i=1}^{q} \bar{\beta}^{-i}} \quad \Leftrightarrow \quad 1 - \alpha = \frac{\bar{\beta}^q (1-\bar{\beta})}{1-\bar{\beta}^q}, \tag{5.10}$$

and, similarly,

$$\bar{\beta} = 1 - \frac{1}{\sum_{i=1}^{p} \alpha^{-i}} \quad \Leftrightarrow \quad 1 - \bar{\beta} = \frac{\alpha^p (1-\alpha)}{1-\alpha^p}. \tag{5.11}$$

For all integer $p \ge 2$ and $x \in [0,1]$, we let

$$f_p(x) = 1 - \frac{1}{\sum_{i=1}^{p} x^{-i}} = 1 - \frac{x^p(1-x)}{1-x^p}.$$

With this notation, we have $\alpha = f_q(\bar{\beta})$ and $\bar{\beta} = f_p(\alpha)$ (and similarly for $\bar{\alpha}$ and β), and thus

$$\alpha = f_q \circ f_p(\alpha) \quad \text{and} \quad \bar{\alpha} = f_q \circ f_p(\bar{\alpha}).$$

We now show that f_p is a contraction for all $p \ge 2$, implying that $f_p \circ f_q$ is also a contraction, and thus admits a unique fixed point, which implies $\alpha = \bar{\alpha}$ (and thus $\beta = \bar{\beta}$).

Lemma 5.3. *For all $p \ge 2$, the function $f_p \colon [0,1] \to \mathbb{R}$ defined by*

$$f_p(x) := 1 - \frac{x^p(1-x)}{1-x^p},$$

is a contraction.

Proof. First note that f_p can be extended to a continuous function on $[0,1]$ by setting $f_p(1) = 1 - 1/p$. To prove that f_p is a contraction, we show that there exists $\varepsilon > 0$ such that, for all $x \in [0,1]$, $|f'_p(x)| \le 1 - \varepsilon$. First note that, for all $x \in [0,1)$,

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

For all $\varepsilon > 0$, we have that

$$\begin{aligned} f'_p(x) \le 1 - \varepsilon \;&\Leftrightarrow\; x^{2p} - (p+1)x^p + px^{p-1} \le 1 - \varepsilon - 2(1-\varepsilon)x^p + (1-\varepsilon)x^{2p} \\ &\Leftrightarrow\; 0 \le 1 - \varepsilon - px^{p-1} + (p-1+2\varepsilon)x^p - \varepsilon x^{2p} =: \varphi(x). \end{aligned}$$

To understand $\varphi(x)$ on $[0,1]$, we look at its derivative: for all $x \in [0,1]$,

$$\varphi'(x) = -p(p-1)x^{p-2} + p(p-1+2\varepsilon)x^{p-1} - 2p\varepsilon x^{2p-1} = x^{p-2}\psi(x),$$

where

$$\psi(x) = -p(p-1) + p(p-1+2\varepsilon)x - 2p\varepsilon x^{p+1}.$$

Note that $\psi'(x) = p(p-1+2\varepsilon) - 2p(p+1)\varepsilon x^p$ is non-negative if and only if

$$x^p \le \frac{p(p-1+2\varepsilon)}{2p(p+1)\varepsilon}.$$

For all ε small enough, the right-hand side of this inequality is larger than one (because $p \ge 2$), implying that for such ε, $\psi'(x)$ is non-negative and thus ψ is non-decreasing on $[0,1]$. And thus, for all $x \in [0,1]$, $\psi(x) \le \psi(1) = 0$. Therefore, since $\varphi'(x) = x^{p-2}\psi(x)$, we get that $\varphi'(x) \le 0$ for all $x \in [0,1]$, and thus φ is non-increasing on $[0,1]$. This implies that $\varphi(x) \le \varphi(0) = 1 - \varepsilon$, and thus concludes the proof. ☐

We have thus proved that $\alpha = \bar{\alpha}$ and $\beta = \bar{\beta}$, where we recall that (α, β) is the unique solution of $\alpha = f_q(\beta)$ and $\beta = f_p(\alpha)$ in $(0,1)^2$, i.e.

$$\begin{cases} \alpha = 1 - \frac{\beta^q(1-\beta)}{1-\beta^q} \\ \beta = 1 - \frac{\alpha^p(1-\alpha)}{1-\alpha^p}. \end{cases} \tag{5.12}$$

It only remains to show that (α, β) is also a solution of (1.1), and that it is the unique solution of (1.1) on $(0,1)^2$. Since (α, β) is a solution of (5.12), we get that

$$(1-\alpha)(1-\beta^q) = \beta^q(1-\beta) \Rightarrow 1-\alpha = \beta^q(2-\alpha-\beta),$$

and, similarly,

$$(1-\beta)(1-\alpha^p) = \alpha^p(1-\alpha) \Rightarrow 1-\beta = \alpha^p(2-\alpha-\beta).$$

This implies

$$2-\alpha-\beta = \frac{1-\beta}{\alpha^p} = \frac{1-\alpha}{\beta^q},$$

and thus (α, β) satisfies the second equation of (1.1). Furthermore,

$$\alpha^p + \beta^q = \frac{1-\beta}{2-\alpha=\beta} + \frac{1-\alpha}{2-\alpha-\beta} = 1,$$

implying that (α, β) is solution of (1.1).

To prove that (1.1) has a unique solution on $(0,1)^2$, we show that any solution of (1.1) on $(0,1)^2$ is also a solution of (5.12) (since the latter has a unique solution on $(0,1)^2$, this concludes the proof). Indeed, if $(\alpha, \beta) \in (0,1)^2$ is a solution of (1.1), then

$$1-\alpha = \frac{\beta^q(1-\beta)}{\alpha^p} = \frac{\beta^q(1-\beta)}{1-\beta^q},$$

which implies the first equation of (5.12). The second equation of (5.12) can obtained similarly. Therefore, if (α, β) is a solution of (1.1), then it is also a solution of (5.12), which concludes the proof of Theorem 1.5.

6 The lozenge: proof of Proposition 1.6

First recall that, by Proposition 2.7 and the specificities of the lozenge graph, if we let $\hat{\mathbf{W}}(n) = \frac{\mathbf{W}(n)}{n+2}$ $(\forall n \geq 0)$, then, for all $n \geq 0$, $\hat{\mathbf{W}}(n) \in \mathcal{E}'$, where

$$\mathcal{E}' := \{\boldsymbol{w} = (w_1, w_2, w_3, w_4, w_5) \in \mathcal{E} \colon w_2 + w_5 = 1, w_1 + w_4 \geq 1, w_2 \leq w_1 + w_3, w_5 \leq w_3 + w_4\},$$

with $\mathcal{E}$ as defined in (2.5). The first condition $(w_2 + w_5 = 1)$ is satisfied by $\hat{\mathbf{W}}(n)$ because, by definition of the model, each ant reinforces either edge 2 or edge 5 but not both. The second condition $(w_1 + w_4 \geq 1)$ is redundant with the fact that $\boldsymbol{w} \in \mathcal{E}$ (it is the same condition as $\pi_{\boldsymbol{w}}(N) \geq 1$). The third condition $(w_2 \leq w_1 + w_3)$ holds because each ant that reinforces edge 2 also reinforces either edge 1 or edge 3. The fourth condition is the symmetric of the third one.

Remark. As in Section 4, it is convenient to define $\hat{\mathbf{W}}(n)$ as $\mathbf{W}(n)/(n+2)$ rather than as $\mathbf{W}(n)/(n+1)$. This is because with the former definition, we have $\hat{W}_2(n) + \hat{W}_5(n) = 1$ for all $n \geq 0$, which is not true with the latter definition (although it holds asymptotically as $n \to +\infty$).

Moreover, for all $n \geq 0$,

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \frac{1}{n+3}\big(F(\hat{\mathbf{W}}(n)) + \xi_{n+1}\big),$$

with ξ_{n+1} some martingale difference, and where $F_i(w) = p_i(w) - w_i$, with $p_i(w) = \mathbb{P}(e_i \in \gamma_{n+1} \mid \hat{\mathbf{W}}(n) = w)$ (note that this probability does not depend on n).

The first step in the proof of Proposition 1.6 is to compute these probabilities $p_i(\boldsymbol{w})$, for $1 \leq i \leq 5$. A straightforward calculation, which we carry out in Section A, shows that for all $\boldsymbol{w} \in \mathcal{E}'$,

$$\begin{aligned}
p_1(\boldsymbol{w}) &= \frac{w_1}{w_1+w_4} + \frac{w_4}{w_1+w_4} \cdot \frac{\frac{w_1}{w_3+w_4+w_5}\left(\frac{w_4}{w_1+w_4} + \frac{w_3}{w_1+w_2+w_3}\right)}{1 - \frac{w_4^2}{(w_1+w_4)(w_3+w_4+w_5)} - \frac{w_3^2}{(w_1+w_2+w_3)(w_3+w_4+w_5)}} \\
p_2(\boldsymbol{w}) &= \frac{w_2(w_1(w_3+w_4+w_5)+w_3w_4)}{(w_1+w_4)(w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4})} \\
p_3(\boldsymbol{w}) &= \frac{w_3\left(\frac{w_1}{w_1+w_2+w_3} + \frac{w_4}{w_3+w_4+w_5}\right)}{w_1+w_4-\left(\frac{w_1^2}{w_1+w_2+w_3} + \frac{w_4^2}{w_3+w_4+w_5}\right)}.
\end{aligned} \tag{6.1}$$

By symmetry, we also have $p_4(\boldsymbol{w}) = p_1(w_4, w_5, w_3, w_1, w_2)$ and $p_5(\boldsymbol{w}) = p_2(w_4, w_5, w_3, w_1, w_2)$. Furthermore, since, by definition of the model, each ant reinforces either edge 2, or edge 5, but not both, we have $p_2(\boldsymbol{w}) = 1 - p_5(\boldsymbol{w})$. Note also that, for all $\boldsymbol{w} \in \mathcal{E}'$,

$$F_2(\boldsymbol{w}) = p_2(\boldsymbol{w}) - w_2 = \frac{w_2w_5(\frac{w_1}{w_1+w_4} - w_2)}{w_3 + w_2w_5 + \frac{w_1w_4}{w_1+w_4}}, \tag{6.2}$$

and

$$p_3(\boldsymbol{w}) = \frac{w_3(\alpha+\beta)}{\alpha(w_3+w_2)+\beta(w_3+w_5)}, \tag{6.3}$$

with

$$\alpha := \frac{w_1}{w_1+w_2+w_3}, \qquad \text{and} \qquad \beta := \frac{w_4}{w_3+w_4+w_5}, \tag{6.4}$$

where by convention we set $\alpha = 0$ when $w_1 = 0$, and similarly $\beta = 0$, when $w_4 = 0$.

The second step is the following fact.

Lemma 6.1. *Almost surely* $\liminf_{n\to+\infty} \frac{W_1(n)}{n} > 0$*, and by symmetry* $\liminf_{n\to+\infty} \frac{W_4(n)}{n} > 0$ *almost surely.*

Proof. Note that for all $\boldsymbol{w} \in \mathcal{E}'$,

$$\frac{p_1(\boldsymbol{w})}{w_1} \geq \frac{1}{w_1+w_4} + \frac{w_4^2}{w_1+w_4} \cdot \frac{1}{(w_3+w_4+w_5)(w_1+w_4) - w_4^2}.$$

When $w_1 \to 0$, we have $w_4 \to 1$ because $1 - w_1 \leq w_4 \leq 1$ for all $\boldsymbol{w} \in \mathcal{E}'$. Using in addition the fact that $w_3 + w_5 \leq 2$ for all $\boldsymbol{w} \in \mathcal{E}'$, we get that

$$\liminf_{w_1\to 0} \frac{p_1(\boldsymbol{w})}{w_1} \geq \frac{3}{2},$$

and then the result follows from Corollary 2.2. □

We next prove the following result.

Lemma 6.2. *For all $\boldsymbol{w} \in \mathcal{E}'$, one has*

(i) *If $w_2 < \frac{w_1}{w_1+w_4}$, then $F_2(\boldsymbol{w}) > 0$, and if $w_2 > \frac{w_1}{w_1+w_4}$, then $F_2(\boldsymbol{w}) < 0$.*

(ii) *If $w_2 \geq \frac{w_1}{w_1+w_4}$, and $0 < w_1 < w_4$, then $F_1(\boldsymbol{w})w_4 - F_4(\boldsymbol{w})w_1 > 0$. Likewise, if $w_2 \leq \frac{w_1}{w_1+w_4}$, and $w_4 < w_1 < 1$, then $F_1(\boldsymbol{w})w_4 - F_4(\boldsymbol{w})w_1 < 0$.*

Proof. The first claim follows directly from (6.2). For the second claim, note that if $w_2 \geq \frac{w_1}{w_1+w_4}$, then with the notation of (6.4), one has $\lambda_1 \leq \frac{w_1}{w_1+\frac{w_1}{w_1+w_4}+w_3}$, and if in addition $w_1 < w_4$, we get

$$\lambda_1 \leq \frac{w_1}{w_1+\frac{w_1}{w_1+w_4}+w_3} < \frac{w_4}{w_4+\frac{w_4}{w_1+w_4}+w_3} \leq \lambda_4,$$

since $w_2 \geq \frac{w_1}{w_1+w_4}$ is equivalent to $w_5 \leq \frac{w_4}{w_1+w_4}$ (using that for $w \in \mathcal{E}$, $w_5 = 1 - w_2$). Then, we get using (6.1), and again $\lambda_4 \geq \lambda_1$, and $w_4 > w_1$,

$$\begin{aligned} F_1(\boldsymbol{w})w_4 - F_4(\boldsymbol{w})w_1 &= p_1(\boldsymbol{w})w_4 - p_4(\boldsymbol{w})w_1 \\ &= \frac{w_1w_4}{w_1+w_4}\left\{\frac{\frac{w_4\lambda_4}{w_1+w_4}+\frac{w_3w_4}{(w_1+w_2+w_3)(w_3+w_4+w_5)}}{1-\frac{w_4\lambda_4}{w_1+w_4}-\frac{w_3^2}{(w_1+w_2+w_3)(w_3+w_4+w_5)}} - \frac{\frac{w_1\lambda_1}{w_1+w_4}+\frac{w_3w_1}{(w_1+w_2+w_3)(w_3+w_4+w_5)}}{1-\frac{w_1\lambda_1}{w_1+w_4}-\frac{w_3^2}{(w_1+w_2+w_3)(w_3+w_4+w_5)}}\right\} > 0, \end{aligned}$$

proving the first statement of (ii). The second statement follows from similar arguments. □

As a corollary we get the following:

Lemma 6.3. *Let $t \mapsto \boldsymbol{\Phi}(t)$ be a solution of the equation $\dot{\boldsymbol{y}} = F(\boldsymbol{y})$, starting from some point $\boldsymbol{w} \in \mathcal{U} := \{\boldsymbol{w} \in \mathcal{E}' : w_1w_4 \neq 0\}$. Then $\lim_{t\to\infty} \boldsymbol{\Phi}(t) = (w^*, 1/2, 1/2, w^*, 1/2)$, where w^* is the unique solution in $[0,1]$ of the equation $2x^3 + 4x^2 - 2x - \frac{3}{2} = 0$.*

Proof. We first show that $\boldsymbol{\Phi}(t)$ converges to the set $\mathcal{H} := \mathcal{E}' \cap \{w_2 = \frac{w_1}{w_1+w_4} = 1/2\}$. Denote by $(\Phi_i(t))_{i=1,\dots,5}$ the coordinates of the vector $\boldsymbol{\Phi}(t)$, and let

$$u(t) = \Phi_2(t) - \frac{1}{2}, \quad \text{and} \quad v(t) = \frac{\Phi_1(t)}{\Phi_1(t)+\Phi_4(t)} - \frac{1}{2}.$$

By definition, taking the derivative along the flow, we get

$$u'(t) = F_2(\boldsymbol{\Phi}(t)), \quad \text{and} \quad v'(t) = \frac{F_1(\boldsymbol{\Phi}(t)) \cdot \Phi_4(t) - F_4(\boldsymbol{\Phi}(t)) \cdot \Phi_1(t)}{(\Phi_1(t)+\Phi_4(t))^2}.$$

Our aim is to show that $h(t) := \max(|u(t)|, |v(t)|)$ is a Lyapunov function, i.e. that it is decreasing, for all t smaller than the (possibly infinite) time when it reaches 0, and that it converges to 0. To see this, first note that if at some time t, one has $0 \leq v(t) < u(t)$, then by Lemma 6.2(*i*), $h'(t) = u'(t) = F_2(\boldsymbol{\Phi}(t)) < 0$. By symmetry, if $u(t) < v(t) \leq 0$, then $h'(t) = -u'(t) < 0$. Second, note that, if $v(t) < 0 \leq u(t)$, then by Lemma 6.2, we have $u'(t) < 0$ and $v'(t) > 0$, which entails that h is decreasing in a neighborhood of t since both its right and left derivatives are negative at this time (it is differentiable if $|u(t)| \neq |v(t)|$). Symmetrically, the same holds if $u(t) < 0 \leq v(t)$. Finally when $u(t) = v(t) \neq 0$, we see that $u'(t) = 0$ while $v'(t) \neq 0$, so that in a neighborhood of t, $u(t) \neq v(t)$, and thus by the previous argument h is again decreasing in a neighborhood of t. As a consequence, h is decreasing up to the (possibly infinite) time when it reaches 0, and thus converges. Note also that the previous arguments show that h has a negative right-derivative at any non-zero value, which implies that its only possible limit is zero. This indeed implies that $\boldsymbol{\Phi}(t)$ converges to the set $\mathcal{H} := \mathcal{E}' \cap \{w_2 = \frac{w_1}{w_1+w_4} = 1/2\}$, as claimed.

We now prove that $\boldsymbol{\Phi}(t)$ converges to the set $\mathcal{H}' := \mathcal{H} \cap \{w_3 = 1/2\}$. Indeed, observe that for any $\boldsymbol{w} \in \mathcal{H}$,

$$F_3(\boldsymbol{w}) = \frac{w_3}{w_3 + 1/2} - w_3,$$

thus $F_3(\boldsymbol{w}) > 0$, if $w_3 < 1/2$, and $F_3(\boldsymbol{w}) < 0$ if $w_3 > 1/2$. Since F_3 is continuous and $\mathcal{E}$ is compact, F_3 is also positive in a neighborhood of $\mathcal{H}\cap\{w_3 \le 1/2-\varepsilon\}$, and negative in a neighborhood of $\mathcal{H}\cap\{w_3 \ge 1/2+\varepsilon\}$, for any fixed $\varepsilon > 0$. Since we also know that $\boldsymbol{\Phi}(t)$ converges to $\mathcal{H}$, it follows that it converges to $\mathcal{H} \cap \{1/2-\varepsilon \le w_3 \le 1/2+\varepsilon\}$, for any $\varepsilon > 0$. In other words it converges well to $\mathcal{H}'$, proving the claim.

Finally, note that for any $\boldsymbol{w} \in \mathcal{H}'$, one has

$$F_1(\boldsymbol{w}) = \frac{1}{2} + \frac{1}{2} \cdot \frac{\frac{w_1}{1+w_1}\left(\frac{1}{2} + \frac{1/2}{1+w_1}\right)}{1 - \frac{1}{2}\frac{w_1}{1+w_1} - \frac{1/4}{(1+w_1)^2}} - w_1 = \frac{1}{2} + \frac{w_1(2+w_1)}{2w_1^2 + 6w_1 + 3} - w_1.$$

Then one can check that $F_1(\boldsymbol{w}) > 0$ if and only if $f(w_1) > 0$ where, for all $x \in \mathbb{R}$,

$$f(x) = -2x^3 - 4x^2 + 2x + \frac{3}{2}.$$

Note that f is a polynomial of degree 3, it thus has at most three zeros in $\mathbb{R}$. One can check that f' is positive on $((-2-\sqrt{7})/3, (-2+\sqrt{7})/3)$ and non-positive on the complement of this set. Thus, on $[0,1]$, f is non-decreasing on $[0, (-2+\sqrt{7})/3]$ and non-increasing on $[(-2+\sqrt{7})/3, 1]$. Since $f(0) = 3/2 > 0$ and $f(1) = -5/2$, we get that there exists a unique solution to $f(x) = 0$ on $[0,1]$, which we call w^*. Moreover, $f(x) > 0$ for all $x \in [0, w^*)$ and $f(x) < 0$ for all $x \in (w^*, 1]$. The conclusion follows, using again continuity of F_1 and compactness of $\mathcal{H}'$, as above. □

The proof of Proposition 1.6 now follows from Corollary 2.6. Indeed, by Lemma 6.1, the limiting set $L(\hat{\boldsymbol{W}})$ of the stochastic approximation $(\hat{\boldsymbol{W}}(n))_{n\ge 0}$ is contained in the set $\mathcal{U}$, which was defined in Lemma 6.3. Then Lemma 6.3 and Corollary 2.6 imply that $L(\hat{\boldsymbol{W}}) = \{(w^*, 1/2, 1/2, w^*, 1/2)\}$, as wanted.

A Calculating F in the lozenge case: proof of (6.1)

We use the same notation as in Section 6. To prove (6.1), we use the electrical networks method (see, e.g. [LL10]). We start by calculating $p_2(\boldsymbol{w})$: this is the probability that a random walker on the graph with weights $\boldsymbol{w} = (w_i)_{1\le i\le 5}$, starting from N, crosses Edge 2 before crossing Edge 5. This is equal to the probability that a walker starting from N reaches F_2 before F_5 on the weighted graph of Figure 6. We decompose $p_2(\boldsymbol{w})$ according to the first step of the walker:

$$p_2(\boldsymbol{w}) = \frac{w_1}{w_1+w_4} \cdot p_{22}(\boldsymbol{w}) + \frac{w_4}{w_1+w_4} \cdot p_{25}(\boldsymbol{w}), \tag{A.1}$$

where $p_{22}(\boldsymbol{w})$ (resp. $p_{25}(\boldsymbol{w})$) denotes the probability to reach F_2 before F_5 starting from P_2 (resp. P_5) on the graph of Figure 6. By classical formulas for random walks on weighted graphs (see, e.g. [LL10]),

$$p_{22}(\boldsymbol{w}) = \frac{\mathcal{C}_{P_2F_2}(\boldsymbol{w})}{\mathcal{C}_{P_2F_2}(\boldsymbol{w}) + \mathcal{C}_{P_2F_5}(\boldsymbol{w})},$$

where $\mathcal{C}_{XY}(\boldsymbol{w})$ is the effective conductance between vertices X and Y when the edge weights are given by $\boldsymbol{w}$. By definition of effective conductances, the effective conductance of a single edge is its weight. Thus, $\mathcal{C}_{P_2F_2} = w_2$. Moreover, the conductance of two edges in parallel is the sum of their effective conductances, the effective conductance of two edges in series is the inverse of the sum of the inverses of their effective conductances. Using these formulas, we get (see Figure 6 for details)

$$\mathcal{C}_{P_2F_2}(\boldsymbol{w}) = \frac{\left(w_3 + \frac{w_1w_4}{w_1+w_4}\right)w_5}{w_3 + w_5 + \frac{w_1w_4}{w_1+w_4}}.$$

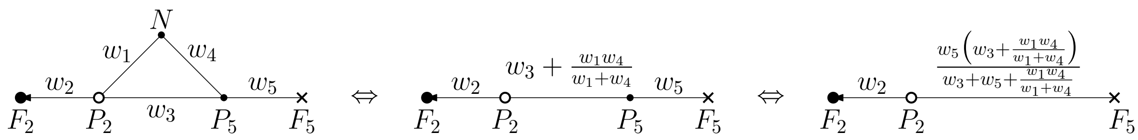

Figure 6: Notation for the proof of (6.1), and calculation of $p_{22}(w)$, the probability of reaching F_2 before F_5 starting from P_2.

We thus get

$$p_{22}(\boldsymbol{w}) = \frac{w_2}{w_2 + \frac{\left(w_3+\frac{w_1w_4}{w_1+w_4}\right)w_5}{w_3+w_5+\frac{w_1w_4}{w_1+w_4}}} = \frac{w_2\left(w_3+w_5+\frac{w_1w_4}{w_1+w_4}\right)}{w_2\left(w_3+w_5+\frac{w_1w_4}{w_1+w_4}\right)+\left(w_3+\frac{w_1w_4}{w_1+w_4}\right)w_5} = \frac{w_2\left(w_3+w_5+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}},$$

because $w_2 + w_5 = 1$ for all $\boldsymbol{w} \in \mathcal{E}'$. By symmetry,

$$p_{25}(\boldsymbol{w}) = 1 - \frac{w_5\left(w_2+w_3+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}} = \frac{(1-w_5)\left(w_3+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}} = \frac{w_2\left(w_3+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}}.$$

Thus, (A.1) becomes

$$\begin{aligned} p_2(\boldsymbol{w}) &= \frac{w_1}{w_1+w_4}\cdot\frac{w_2\left(w_3+w_5+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}} + \frac{w_4}{w_1+w_4}\cdot\frac{w_2\left(w_3+\frac{w_1w_4}{w_1+w_4}\right)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}} \\ &= \frac{w_2(w_1(w_3+w_4+w_5)+w_3w_4)}{w_3+w_2w_5+\frac{w_1w_4}{w_1+w_4}}, \end{aligned}$$

as claimed.

We now calculate $p_3(\boldsymbol{w})$: we decompose on the first step of the random walker to get

$$p_3(\boldsymbol{w}) = \frac{w_1}{w_1+w_4}\cdot p_{32}(\boldsymbol{w}) + \frac{w_4}{w_1+w_4}\cdot p_{35}(\boldsymbol{w}), \tag{A.2}$$

where $p_{32}(\boldsymbol{w})$ (resp. $p_{35}(\boldsymbol{w})$) is the probability to cross edge 3 before reaching F starting from P_2 (resp. P_5), when the edge weights are given by $\boldsymbol{w}$. Decomposing over the first weight of a random walker starting at P_2, we get

$$p_{32}(\boldsymbol{w}) = \frac{w_1}{w_1+w_2+w_3}\cdot p_3(\boldsymbol{w}) + \frac{w_3}{w_1+w_2+w_3},$$

and similarly for $p_{35}(\boldsymbol{w})$. Using this in (A.2), we get

$$\begin{aligned} &p_3(\boldsymbol{w}) \\ &= \frac{w_1}{w_1+w_4}\left(\frac{w_1}{w_1+w_2+w_3}\cdot p_3(\boldsymbol{w}) + \frac{w_3}{w_1+w_2+w_3}\right) + \frac{w_1}{w_1+w_4}\left(\frac{w_4}{w_3+w_4+w_5}\cdot p_3(\boldsymbol{w}) + \frac{w_3}{w_3+w_4+w_5}\right), \end{aligned}$$

which implies

$$\begin{aligned} &\left(1 - \frac{w_1^2}{(w_1+w_4)(w_1+w_2+w_3)} - \frac{w_4^2}{(w_1+w_4)(w_3+w_4+w_5)}\right)p_3(\boldsymbol{w}) \\ &\qquad = \frac{w_3}{w_1+w_4}\left(\frac{w_1}{w_1+w_2+w_3} + \frac{w_4}{w_3+w_4+w_5}\right). \end{aligned}$$

This indeed gives the formula for $p_3(\boldsymbol{w})$ announced in (6.1).

Finally, we show how to calculate $p_1(\boldsymbol{w})$: again, we decompose according to the first step of the walker:

$$p_1(\boldsymbol{w}) = \frac{w_1}{w_1 + w_4} + \frac{w_4}{w_1 + w_4} \cdot p_{15}(\boldsymbol{w}), \tag{A.3}$$

where $p_{15}(\boldsymbol{w})$ is the probability to cross edge 1 before reaching F starting from P_5. Decomposing according to the first step again, we get

$$\begin{aligned} p_{15}(\boldsymbol{w}) &= \frac{w_4}{w_3 + w_4 + w_5} \cdot p_1(\boldsymbol{w}) + \frac{w_3}{w_3 + w_4 + w_5} \cdot p_{12}(\boldsymbol{w}) \\ &= \frac{w_4}{w_3 + w_4 + w_5} \left(\frac{w_1}{w_1 + w_4} + \frac{w_4}{w_1 + w_4} \cdot p_{15}(\boldsymbol{w}) \right) + \frac{w_3}{w_3 + w_4 + w_5} \cdot p_{12}(\boldsymbol{w}) \end{aligned} \tag{A.4}$$

where $p_{12}(\boldsymbol{w})$ is the probability to cross edge 1 before reaching F starting from P_2. We have used (A.3) in the second equality. Finally, we have

$$p_{12}(\boldsymbol{w}) = \frac{w_1}{w_1 + w_2 + w_3} + \frac{w_3}{w_1 + w_2 + w_3} \cdot p_{15}(\boldsymbol{w}). \tag{A.5}$$

Using (A.5) in (A.4) we get

$$p_{15}(\boldsymbol{w}) = \frac{\frac{w_1}{w_1+w_4} \cdot \frac{w_4}{w_3+w_4+w_5} + \frac{w_1}{w_1+w_2+w_3} \cdot \frac{w_3}{w_3+w_4+w_5}}{1 - \frac{w_4^2}{(w_1+w_4)(w_3+w_4+w_5)} - \frac{w_3^2}{(w_1+w_2+w_3)(w_3+w_4+w_5)}},$$

and we can then use in (A.3) to get the formula for $p_1(\boldsymbol{w})$ announced in (6.1).

References

[BBCL15] Michel Benaïm, Itai Benjamini, Jun Chen, and Yuri Lima. A generalized Pólya's urn with graph based interactions. *Random Structures & Algorithms*, 46(4):614–634, 2015. 2, 3

[Ben99] Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Seminaire de probabilités XXXIII*, pages 1–68. Springer, 1999. 3, 6, 7

[CH21] Yannick Couzinié and Christian Hirsch. Weakly reinforced Pólya urns on countable networks. *Electronic Communications in Probability*, 26:1–10, 2021. 3

[CL14] Jun Chen and Cyrille Lucas. A generalized Pólya's urn with graph based interactions: convergence at linearity. *Electronic Communications in Probability*, 19, 2014. 3

[DAGP90] Jean-Louis Deneubourg, Serge Aron, Simon Goss, and Jacques M. Pasteels. The self-organizing exploratory pattern of the argentine ant. *Journal of insect behavior*, 3(2):159–168, 1990. 1

[DS04] Marco Dorigo and Thomas Stützle. *Ant colony optimization*. MIT Press, 2004. 1, 2

[EFR19] Dirk Erhard, Tertuliano Franco, and Guilherme Reis. The directed edge reinforced random walk: The ant mill phenomenon. https://arxiv.org/abs/1911.07295, 2019. 2

[GADP89] Simon Goss, Serge Aron, Jean-Louis Deneubourg, and Jacques Marie Pasteels. Self-organized shortcuts in the argentine ant. *Naturwissenschaften*, 76(12):579–581, 1989. 1

[Ham20] Yassine Hamdi. A model of ants' search for food by reinforced walks on graphs (UG internship report). https://gitlab.com/yassine_hamdi/a-model-of-ants-search-for-food-by-reinforced-walks-on-graphs-public, 2020. 5

[HHK20] Christian Hirsch, Mark Holmes, and Victor Kleptsyn. Warm percolation on a regular tree in the strong reinforcement regime. https://arxiv.org/abs/2009.07682, 2020. 3

[HJ04] Ben M. Hambly and Jonathan Jordan. A random hierarchical lattice: the series-parallel graph and its properties. *Advances in applied probability*, pages 824–838, 2004. 4

[HK] Mark Holmes and Victor Kleptsyn. Infinite WARM graphs II: critical regime. 3

[Jan04] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110(2):177–245, 2004. 3

[KMS] Daniel Kious, Cécile Mailler, and Bruno Schapira. Finding geodesics on graphs using reinforcement learning. https://arxiv.org/abs/2010.04820. 1, 2, 4, 5, 8

[LGR18] Line C. Le Goff and Olivier Raimond. Vertex reinforced non-backtracking random walks: an example of path formation. *Electronic Journal of Probability*, 23, 2018. 2

[Lim16] Yuri Lima. Graph-based Pólya's urn: completion of the linear case. *Stochastics and Dynamics*, 16(02):1660007, 2016. 3

[LL10] Gregory F. Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123. Cambridge University Press, 2010. 12, 24

[LP05] Russell Lyons and Yuval Peres. *Probability on trees and networks.* Cambridge University Press, 2005. 9, 10

[Mar17] Andreï A. Markov. Sur quelques formules limites du calcul des probabilités (in Russian). *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, 11(3):177–186, 1917. 3

[MJT$^+$13] Qi Ma, Anders Johansson, Atsushi Tero, Toshiyuki Nakagaki, and David J.T. Sumpter. Current-reinforced random walks for constructing transport networks. *Journal of the Royal Society Interface*, 10(80):20120864, 2013. 1

[Pem07] Robin Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, pages 1–79, 2007. 3, 6, 7

[VDHHKR16] Remco Van Der Hofstad, Mark Holmes, Alexey Kuznetsov, and Wioletta Ruszel. Strongly reinforced Pólya urns with graph-based competition. *The Annals of Applied Probability*, 26(4):2494–2539, 2016. 2, 3