# Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk

Roberto Imbuzeiro Oliveira*

July 19, 2010

**Abstract**

We prove an upper bound for the $\epsilon$-mixing time of the symmetric exclusion process on any graph $G$, with any feasible number of particles. Our estimate is proportional to $\mathsf{T}_{\mathsf{RW}(G)} \ln(|V|/\epsilon)$, where $|V|$ is the number of vertices in $G$ and $\mathsf{T}_{\mathsf{RW}(G)}$ is the 1/4-mixing time of the corresponding single-particle random walk. This bound implies new results for symmetric exclusion on expanders, percolation clusters, the giant component of the Erdös-Rényi random graph and Poisson point processes in $\mathbb{R}^d$. Our technical tools include a variant of Morris' chameleon process.

## 1 Introduction

The *symmetric exclusion process* is a continuous-time Markov Chain defined on a *weighted graph* $G = (V, E, \{w_e\}_{e\in E})$, where $V$ is a set of vertices, $E$ is a set of edges and to each $e \in E$ we assign a positive weight $w_e > 0$. For $k \leq |V|$, $k$-particle symmetric exclusion on $G$ has the following informal description.

> $\mathsf{EX}(k, G)$: Start with $k$ indistinguishable particles placed on distinct vertices of $V$. Each particle moves independently according to the symmetric transition rates given by the edge weights, except that moves to occupied sites are suppressed.

This is one of the most basic and best studied processes in the literature on Interacting Particle Systems [15, 16]. Literally hundreds of papers have been written on this process, but most of these results apply only to restricted classes of infinite graphs, such as the lattices $\mathbb{Z}^d$.

Exclusion processes over finite graphs have also been a testbed for the quantitative analysis of finite Markov chains. Coupling [2], comparison arguments [9], the martingale method for log-Sobolev inequalities [12, 23], and variants of the evolving sets technology

*IMPA, Rio de Janeiro, RJ, Brazil, 22430-040. rimfo@impa.br

[20, 21] have been variously applied to this process. Sharp results are known for some special cases, such as the complete graph [12] and discrete tori $(\mathbb{Z}/L\mathbb{Z})^d$ [20, 23].

In this paper we consider $\mathsf{EX}(k,G)$ over an arbitrary finite graph and bound its mixing time in terms of the corresponding single-particle random walk, which we denote by $\mathsf{RW}(G)$. Our result is very general, but we will see that it matches or improves upon most previously known mixing results for $\mathsf{EX}(k,G)$. We will also argue in this introduction that the kind of result presented here is of conceptual interest.

## 1.1 The main result (and why it is interesting)

We recall what mixing times are before we state our main result. An irreducible continuous-time Markov chain $Q$ on a finite set $S$, with transition probabilities $\{q_t(s,s')\}_{s,s'\in S, t\geq 0}$, always has a unique stationary (equilibrium) distribution $\pi$, to which the transition probabilities converge. The $\epsilon$-mixing time $\mathsf{T}_Q(\epsilon)$ quantifies the speed of this convergence:

$$\mathsf{T}_Q(\epsilon) \equiv \inf\left\{ t \geq 0 \,:\, \max_{s\in S} d_{\mathrm{TV}}(q_t(s,\cdot),\pi) \leq \epsilon \right\}, \tag{1.1.1}$$

where $d_{\mathrm{TV}}$ is the total-variation distance for probability measures over $S$ (cf. (2.2.1)). The 1/4 mixing time $\mathsf{T}_Q(1/4)$ will also be called *the* mixing time of $Q$. Our main result is:

**Theorem 1.1.1 (Main result; proven in Section 6.1)** *There exists a universal constant $C > 0$ such for all $\epsilon \in (0,1)$, all connected weighted graphs $G = (V, E, \{w_e\}_{e\in E})$ with $|V| \geq 2$ and all $k \in \{1,\ldots,|V|-1\}$,*

$$\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq C\, \ln(|V|/\epsilon)\, \mathsf{T}_{\mathsf{RW}(G)}(1/4).$$

While this might not be obvious at first sight, our bound follows quite naturally if one *assumes* (heuristically) that the mixing time of $\mathsf{EX}(k,G)$ is not much larger than that of $k$ independent random walks on $G$, a process we denote by $\mathsf{RW}(k,G)$ in what follows.

**[Heuristic assumption]** $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq C_0\, \mathsf{T}_{\mathsf{RW}(k,G)}(\epsilon), C_0 > 0$ universal.

This assumption is rather strong, not at all obvious and, as far as we can tell, beyond the reach of present techniques. However, it seems at least plausible, given that $\mathsf{RW}(k,G)$ and $\mathsf{EX}(k,G)$ are similar.

It can be shown that $\mathsf{T}_{\mathsf{RW}(k,G)}(\epsilon) = \Theta\left(\mathsf{T}_{\mathsf{RW}(G)}(\epsilon/k)\right)$ as $\epsilon/k \to 0$, thus our assumption is equivalent to:

**[Heurisitic assumption]'** $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq C_1\, \mathsf{T}_{\mathsf{RW}(G)}(\epsilon/k), C_1 > 0$ universal..

Recall the general inequality "$\mathsf{T}_{\mathsf{RW}(G)}(\delta) \leq C_2\, \ln(1/\delta)\, \mathsf{T}_{\mathsf{RW}(G)}(1/4)$", with $C_2 > 0$ universal, which is valid for any $0 < \delta < 1/2$ [2]. Applying this to our assumption, we obtain:

**[Heuristic conclusion]** $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq C_3\, \ln(k/\epsilon)\, \mathsf{T}_{\mathsf{RW}(G)}(1/4), C_3 > 0$ universal.

Theorem 1.1.1 coincides with this for $k > |V|^c$, $c > 0$ constant; whereas for other $k$ it is a strictly weaker result.

We emphasize that what we just presented is *not* a rigorous proof of Theorem 1.1.1, since we offer no good grounds for our heuristic assumption. What is interesting is that the Theorem does give an *a posteriori* justification for a weakened form of the assumption. We note that the bound "$\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq C\mathsf{T}_{\mathsf{RW}(G)}(1/4)\ln(k/\epsilon)$" is tight up to constant factors for some $G$ (e.g. discrete tori $(\mathbb{Z}/L\mathbb{Z})^d$, $d$ fixed [20]); therefore, in some sense Theorem 1.1.1 is quite close to the best that one might reasonably hope for.

Many other complex Markov chains are built from simpler processes that interact; examples appear in eg. [2, 8, 18]. Given our main result, it seems reasonable that, at least in some cases, the mixing time of these complex processes may be bounded in terms of their constituent parts. Some of the techniques we use to prove Theorem 1.1.1 are very specific to $\mathsf{EX}(k,G)$, but it may be that some of the same ideas will turn out to be useful in other cases.

## 1.2 Connections with Aldous' conjecture

Another motivation for our paper is a conjecture of Aldous' for the *interchange process*, which was recently proved in [7]. The *interchange process on $G$ with $k \leq |V|$ particles* can be informally described as follows:

> $\mathsf{IP}(k,G)$: Start with $k$ particles labelled $1, 2, \ldots, k$ placed on distinct vertices of $V$ all remaining vertices (if any) are labelled "empty". For each edge $e$, switch the labels of the endpoints of $e$ at rate $w_e$.

One can obtain $\mathsf{EX}(k,G)$ from $\mathsf{IP}(k,G)$ by "forgetting" the labels of the $k$ particles. In particular, the contraction principle [2] implies that $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq \mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon)$ for all $1 \leq k \leq |V| - 1$ and all $\epsilon \in (0,1)$.

Aldous conjectured – and Caputo et al. recently proved [7] (see also [11]) – that $\mathsf{IP}(k,G)$ and $\mathsf{EX}(k,G)$ always have the same *spectral gap* as $\mathsf{RW}(G)$ (or $\mathsf{RW}(k,G)$). This is a remarkable result, but it does not say much about the *mixing times* of these processes, since the bounds for $\mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon)$ or $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon)$ that can be obtained from the spectral gap are typically very loose.

Theorem 1.1.1 gives tighter relations between these mixing times. In the proof of the Theorem, we will show that the bound claimed for $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon)$ in the Theorem statement in fact holds for $\mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon)$ whenever $k \leq |V|/2$. Our proofs can be adapted to show that:

$$\forall \alpha \in (0,1),\ \exists C_\alpha > 0,\ \forall G,\ \forall k \leq \alpha|V| \ :\ \mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon) \leq C_\alpha\, \mathsf{T}_{\mathsf{RW}(G)}(1/4)\ \ln(|V|/\epsilon).$$

That is, one can get a bound similar to Theorem 1.1.1 also for the interchange process, as long as the fraction of empty sites is bounded away from 0. Unfortunately, this leaves out

the most interesting case of $\mathsf{IP}(|V|, G)$, which is a random walk by random transpositions in the group of permutations of $V$. Fortunately, the restriction on $k$ does not make a difference for the exclusion process.

### 1.3 Applications and comparison with previous results

It is not hard to apply Theorem 1.1.1 to specific examples: all one needs is a bound for the mixing time of simple random walk on the given graph, and $\mathsf{RW}(G)$ is typically much easier to analyse than $\mathsf{EX}(k, G)$. Some applications are discussed in detail in Section 10, where we also compare Theorem 1.1.1 with previous results in the literature. Here is a summary of that section.

- The best previous general mixing time bound relies on comparison arguments, cf. Section 10.1.
- Morris' optimal result for the discrete torus $(\mathbb{Z}/L\mathbb{Z})^d$ [20] is that the mixing time of $\mathsf{EX}(k, G)$ is $\Theta\left(L^2 \ln k\right)$. Our bound equals his up to constant factors for $k \geq L^{cd}$, $c > 0$ some universal constant, and matches the previous best result, which could be derived from the general bounds mentioned above. See Section 10.2 for details.
- On the other hand, the bound for large percolation clusters over $\{1, 2, \ldots, L\}^d$ that we present in Section 10.3 is also $O\left(L^2 \ln L\right)$, and it seems very difficult (if not impossible) to obtain this via the comparison method, or via a direct adaptation of Morris' methods.
- For constant-degree expanders, the comparison bound will always be at least $\Omega\left(\ln^3 |V|\right)$, whereas our bound is $O\left(\ln^2 |V|\right)$. See Section 10.4 for details.
- In Section 10.5 we obtain a $O\left(\ln^3 |V|\right)$ bound for the giant component of a supercritical Erdös-Rényi graph. This matches the best possible result that one could *hope* to obtain via the comparison bound. It is not at all clear that our bound can indeed be matched by the comparison bound. This is discussed in Section 10.5.
- We also obtain results on certain random graphs graphs built on $[0, L]^d \cap \mathcal{P}$, where $\mathcal{P}$ is a Poisson processes over $\mathbb{R}^d$. See Section 10.6 for details.

### 1.4 Organization and key steps of the proof

Our proof of Theorem 1.1.1 will use many elementary facts about coupling, total variation distance and Markov chains, and *Section 2* is devoted to these and other preliminaries. *Section 3* reviews the well-known graphical construction of $\mathsf{RW}(G), \mathsf{IP}(k, G)$ and $\mathsf{EX}(k, G)$.

Many steps in our proof follow from analyzing this construction or related constructions introduced later on. Another fundamental fact, the *negative correlation property* of $\mathsf{EX}(k, G)$, is also recalled in that section.

Actual proofs start in *Section 4*, where we study the trajectories of *pairs of labelled particles*, ie. $\mathsf{IP}(2, G)$. The main result in that section is that $\mathsf{IP}(2, G)$ always has a mixing time comparable to $\mathsf{RW}(G)$.[1]

**Lemma 1.4.1** *For any weighted graph $G$,*

$$\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \leq 20000\, \mathsf{T}_{\mathsf{RW}(G)}(1/4).$$

The proof of this Lemma relies on realizing that there are two classes of graphs. Some $G$ are "easy", in that two independent random walkers are likely to meet by time $O\left(\mathsf{T}_{\mathsf{RW}(G)}(1/4)\right)$ from *any pair* of initial states. In this case, an argument of Aldous and Fill [2] suffices to prove Lemma 1.4.1 (see Proposition 4.2.1).

On the other hand, if $G$ is not easy, then for *most* initial states two independent random walkers are very unlikely to meet by time $\Omega\left(\mathsf{T}_{\mathsf{RW}(G)}(1/4)\right)$ (cf. Proposition 4.3.1). It is not hard to show that $\mathsf{RW}(2, G)$ and $\mathsf{IP}(2, G)$ are similar for a long time, if started from vertices from which meeting takes a long time (Proposition 4.4.1 and Corollary 4.4.1). Since *most* pairs of vertices are like that, $\mathsf{IP}(2, G)$ and $\mathsf{RW}(2, G)$ are similar after a burn-in period which is long enough for a favorable pair of vertices to be reached. This will ultimately lead to Lemma 1.4.1 for non-easy graphs. We note that our argument requires the negative correlation property at a crucial step; see Remark 4.5.1 for details.

The remainder of the proof of Theorem 1.1.1 consists of bootstrapping Lemma 1.4.1 to a larger number of particles. We will eventually show a Lemma that implies the main Theorem, given Lemma 1.4.1:

**Lemma 1.4.2 (Proven in Section 6.2)** *There exists a universal constant $K > 0$ such that for all connected weighted graphs $G = (V, E, \{w_e\}_{e \in E})$ and all $k \in \{1, \ldots, |V|/2\}$,*

$$\mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon) \leq K\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)\, \ln(|V|/\epsilon).$$

There is at least two methods in the literature for moving from pairs of particles to many more particles. Both of them were introduced by Morris [19, 20]. The first one comes from [19] and gives bounds for walks on the symmetric group by random transpositions. This could be in principle applied to $\mathsf{IP}(|V|, G)$, but the method seems to require too much from the process to be useful in our general setting. Even if this difficulty can be surmounted, the bounds given by that method would have a factor of $\ln(|V|)\ln(1/\epsilon)$ where ours has a $\ln(|V|/\epsilon)$ term.

[1] Since the single-particle marginal distributions of $\mathsf{IP}(2, G)$ are given by $\mathsf{RW}(G)$, $\mathsf{T}_{\mathsf{RW}(G)}(1/4) \leq \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$ is immediate from the contraction principle.

Morris' other method was introduced in his study of symmetric exclusion over $(\mathbb{Z}\backslash L\mathbb{Z})^d$ [20]. The so-called *chameleon process* features particles that change color in a way that encodes the conditional distribution of the $k$-th particle in $\mathsf{IP}(k, G)$ given the other $k-1$ particles. It is this method that we will successfully adapt to prove Lemma 1.4.2.

The main idea, both in his construction and in our variant, is that the study of mixing is reduced to the study of pairwise collisions between particles. The analysis for $(\mathbb{Z}\backslash L\mathbb{Z})^d$ is greatly facilitated by the explicit structure of the graph, something that we lack in general. This difficulty will require certain technical modifications of Morris' construction of which we will try to make sense with remarks in our proofs.

Our modified chameleon process is presented in *Section 5*. We prove some basic properties of the process in that section; in particular, we show how the transition probabilities of $\mathsf{IP}(k, G)$ can be expressed by the chameleon process. In *Section 6* we prove Lemma 1.4.2 and Theorem 1.1.1 using many facts about the chameleon process that we will prove later on. The remainder of the proof is sketched in Section 6.3 and carried out in *Sections 7 through 9*. It would be pointless to try to describe these steps now, but we notice that the modifications we introduce in the chameleon process are elucidated only towards the end of the proof.

Applications of the main result and comparisons are performed in Section 10. Section 11 presents some final remarks. An Appendix contains some technical results, including some steps where we follow Morris' original argument quite closely.

# 2 Preliminaries

## 2.1 Basic notation

$\mathbb{N} = \{0, 1, 2, 3, \dots\}$ is the set of non-negative integers and $\mathbb{N}_+ \equiv \mathbb{N}\backslash\{0\}$. For $n \in \mathbb{N}_+$, $[n] \equiv \{i \in \mathbb{N}_+ : i \leq n\} = \{1, \dots, n\}$. If $S$ is a finite set, $|S|$ is the cardinality of $S$. For any $k \in [|S|]$,

$$\binom{S}{k} = \{A \subset S : |A| = k\}$$

is the set of all size-$k$ subsets of $S$ and

$$(S)_k = \{\mathbf{s} = (\mathbf{s}(1), \dots, \mathbf{s}(k)) \in S^k : \forall i, j \in [k], \text{“}i \neq j\text{”} \Rightarrow \text{“}\mathbf{s}(i) \neq \mathbf{s}(j)\text{”}\}$$

is the set of all $k$-tuples of distinct elements in $S$.

**Notational convention 2.1.1** *The elements of $(S)_k$ will always be denoted by boldface letters such as $\mathbf{x}$, with $\mathbf{x}(i)$ denoting the $i$-th coordinate of $\mathbf{x}$.*

Notice that with these symbols,

$$\left|\binom{S}{k}\right| = \binom{|S|}{k}, \ |(S)_k| = (|S|)_k.$$

A *graph* is a couple $H = (V, E)$ where $V \neq \emptyset$ is the set of *vertices* and $E \subset \binom{V}{2}$ is the set of *edges*. For each $e \in E$, the two elements $a, b \in V$ such that $e = \{a, b\}$ are called the *endpoints of* $e$.

A *weighted graph* is a triple $G = (V, E, \{w_e\}_{e \in E})$ where $(V, E)$ is a graph and $w_e > 0$, the *weight of edge* $e$, is positive for each $e \in E$. When a graph $G$ is introduced without explicitly defining the edge weights, we will assume that they are all equal to 1. We will assume throughout this paper that all graphs we consider are connected.

## 2.2 Basic probabilistic concepts

Let $(\Omega, \mathcal{F})$ and $(\Theta, \mathcal{G})$ be measurable spaces and $X : \Omega \to \Theta$ be a random variable, ie. a $\mathcal{F}/\mathcal{G}$-measurable map ($\mathcal{G}$ will usually be implicit in our discussions). A probability measure $\mathbb{P}$ over $(\Omega, \mathcal{F})$ induces a measure $\mathcal{L}[X]$ over $(\Theta, \mathcal{G})$ via $\mathcal{L}[X](G) \equiv \mathbb{P}(X \in G)$, $G \in \mathcal{G}$. $\mathcal{L}[X]$ will be called the *law* or *distribution* of $X$.

Given two probability distributions $\mu, \nu$ over the same finite set $S$, the *total variation distance* between them is given by several equivalent formulas:

$$\begin{aligned} d_{\mathrm{TV}}(\mu, \nu) &\equiv \max_{A \subset S} (\mu(A) - \nu(A)) && (2.2.1) \\ &= \sup_{f : S \to [0,1]} \int f \, d\mu - \int f \, d\nu && (2.2.2) \\ &= \sum_{s \in S} (\mu(s) - \nu(s))_+ && (2.2.3) \\ &= \frac{1}{2} \sum_{s \in S} |\mu(s) - \nu(s)| && (2.2.4) \end{aligned}$$

Another equivalent definition of $d_{\mathrm{TV}}$ is

$$d_{\mathrm{TV}}(\mu, \nu) = \inf \mathbb{P}(X \neq Y),$$

where the infimum is over all pairs $(X, Y)$ of $S^2$-valued random variable with $\mathcal{L}[X] = \mu$ and $\mathcal{L}[Y] = \nu$ (such a pair is called a *coupling* of $(\mu, \nu)$). This implies that for any pair of $S$-valued random variables $X, Y$ defined over the same probability space:

$$d_{\mathrm{TV}}(\mathcal{L}[X], \mathcal{L}[Y]) \leq \mathbb{P}(X \neq Y)$$

We will need the following simple fact: if (for $i = 1, 2$) $\mu_i, \nu_i$ are probability distributions on the finite set $S_i$,

$$d_{\mathrm{TV}}(\mu_1 \times \mu_2, \nu_1 \times \nu_2) \leq d_{\mathrm{TV}}(\mu_1, \nu_1) + d_{\mathrm{TV}}(\mu_2, \nu_2). \tag{2.2.5}$$

We will write $\mathrm{Unif}(S)$ for the *uniform distribution* on a set $S \neq \emptyset$. This is the normalized counting measure on $S$, if $S$ is finite, or normalized Lebesgue measure over $S$, if $S \subset \mathbb{R}^d$.

### 2.3 Markov chains and mixing times

For our purposes it is convenient to define a continous-time Markov chain over a finite set $S$ as a family of processes

$$\{(X_t^s)_{t\geq 0} \,:\, s \in S\}$$

defined on the same probability space, with the following properties:

1. For each $s \in S$, $X_0^s = s$ almost surely;
2. Each $X_t^s$ is a "càdlàg" path over $S$: there exists a divergent sequence

$$\tau_0 = 0 < \tau_1 < \tau_2 < \dots$$

   and a sequence $\{s_i\}_{i\geq 0} \subset S$ with $s_0 = s$ with $X_t^s \equiv s_i$ over each interval $[\tau_i, \tau_{i+1})$.
3. For each $h \geq 0$ and each càdlàg path $(x_u)_{u\geq 0}$ taking values in $S$ (in the sense of 2.),

$$\mathbb{P}\left(X_{t+h}^s = s' \mid X_{t'}^s = x_{t'}, 0 \leq t' \leq t\right) = \mathbb{P}\left(X_h^{x_t} = s'\right) \text{ almost surely.}$$

The last property is the so-called *Markov property*. It also implies that the law of $(X_{t+h}^s)_{h\geq 0}$ equals that of $(X_h^{x_t})_{h\geq 0}$ under the above conditioning. It is well-kwown that any such process is uniquely defined by its transition rates

$$q(s, s') \equiv \lim_{\epsilon \searrow 0} \frac{\mathbb{P}\left(X_\epsilon^s = s'\right)}{\epsilon} \quad ((s, s') \in S^2,\ s \neq s')$$

or equivalenty by its *generator*:

$$Q : f \in \mathbb{R}^S \mapsto Qf(\cdot) \equiv \sum_{s' \in S,\, s' \neq \cdot} q(\cdot, s')(f(s') - f(\cdot)).$$

We will usually make no distinction between a Markov chain and its generator in our notation.

In this paper we will only work with *irreducible chains*, ie. chains for which for all $A \subset V$ with $A \neq \emptyset$, $V \backslash A \neq \emptyset$, there exist $a \in A$, $b \in V \backslash A$ with $q(a, b) > 0$. It is well-known that such Markov chains have a unique *stationary* distribution $\pi$, ie. a distribution such that if $s_*$ is picked according to $\pi$ independently from the $(X_t^s)_{t\geq 0, s\in S}$, then $\mathcal{L}[X_t^{s_*}] = \pi$ for all $t \geq 0$. Moreover,

$$\forall s \in S,\ d_{\mathrm{TV}}(\mathcal{L}[X_t^s], \pi) \to 0 \text{ as } t \to +\infty.$$

The *$\epsilon$-mixing time of $Q$* is thus defined as in the Introduction:

$$\mathsf{T}_Q(\epsilon) \equiv \inf\{t \geq 0 \,:\, \max_{s\in S} d_{\mathrm{TV}}(\mathcal{L}[X_t^s], \pi) \leq \epsilon\} \quad (\epsilon \in (0, 1)).$$

We will often need two elementary facts about Markov chains and their mixing times.

**Proposition 2.3.1 ([13], equation (4.36) in page 55)** *Let $Q$ be a Markov chain on finite state space $S$. Then for all $0 < \epsilon < 1/2$,*[2]

$$\mathsf{T}_Q(\epsilon) \leq \lceil \log_2(1/\epsilon) \rceil \, \mathsf{T}_Q(1/4).$$

**Proposition 2.3.2 ([2], Lemma 7 in Chapter 4)** *Let $Q$ be a Markov chain on finite state space $S$ with symmetric transition rates. Then $\pi$ is uniform over $S$ and moreover:*

$$\mathbb{P}\left(X_t^s = s'\right) \geq \frac{(1-2\epsilon)^2}{|S|},$$

*for all $s, s' \in S$, with the same notation introduced above.*

We also make the following convenient notational convention.

**Notational convention 2.3.1** *By definition, for any càdlag path $(x_t)_{t\geq 0}$ there exists a divergent sequence $t_0 = 0 < t_1 < t_2 < \dots$ with $x_t$ constant over $[t_i, t_{i+1})$ for each $i \geq 0$. For $t > 0$, we define $x_{t\_}$ to be the state of $x_t$ immediately prior to time $t$. That is,*

$$x_{t\_} \equiv \begin{cases} x_{t_{i-1}}, & \text{if } t = t_i \text{ for some } i \geq 1; \\ x_t & \text{otherwise.} \end{cases}$$

*Notice that $x_{t\_} = x_{t-\delta}$ for all $\delta > 0$ sufficiently small.*

# 3 Random walks, exclusion and interchange processes

In this section we formally define the main Markov chains in this paper: $\mathsf{RW}(G), \mathsf{EX}(k, G)$ and $\mathsf{IP}(k, G)$. We also present the standard *graphical construction* for the three processes at the same time, and then discuss the negative correlation property for $\mathsf{EX}(k, G)$.

The material in this section is quite classical: Liggett's books [15, 16] are basic references , and the manuscript by Aldous and Fill [2] contains some additional facts on $\mathsf{IP}(k, G)$ as well as a presentation that is somewhat closer in style to ours.

## 3.1 Definitions

The three processes we are defined in terms of the same *weighted graph* $G = (V, E, \{w_e\}_{e\in E})$ with $V$ finite (cf. Section 2). We will be implicitly assuming that $G$ is connected, in which case one can easily show that the chains defined below are irreducible.

[2] The result in [13] is for discrete-time chains, but the proof trivially extends to continuous time.

*Simple random walk on* $G$, denoted by $\mathsf{RW}(G)$, is the continuous-time Markov chain with state space $V$ and transition rates:

$$q(u,v) \equiv \begin{cases} w_e & \text{if } e = \{u,v\} \in E; \\ 0 & \text{otherwise.} \end{cases} \quad ((u,v) \in (V)_2).$$

If we define the transpositions:

$$f_e : x \in V \mapsto \begin{cases} b & \text{if } x = a \\ a & \text{if } x = b \\ x & \text{otherwise.} \end{cases},$$

we see that $q(u,v) = w_e$ if $v = f_e(u)$ for some (necessarily unique) $e \in E$ and $q(u,v) = 0$ otherwise. It follows the *generator* $Q_{\mathsf{RW}(G)}$ of $\mathsf{RW}(G)$ acts on functions $h \in \mathbb{R}^V$ as follows.

$$Q_{\mathsf{RW}(G)}(h) = \sum_{e \in E} w_e \, (h \circ f_e - h). \tag{3.1.1}$$

We will also consider the process $\mathsf{RW}(k,G)$ that corresponds to $k$ such random walks performed simultaneously and independently. Since the transition rates of this process are symmetric, it follows that the stationary distribution of $\mathsf{RW}(k,G)$ is $\mathrm{Unif}(V^k)$ for all $k$.

*The* $k$*-particle symmetric exclusion process on* $G$, denoted by $\mathsf{EX}(k,G)$, is the continuous-time Markov chain with state space $\binom{V}{k}$ and transition rates:

$$q^{\{k\}}(A,B) \equiv \begin{cases} w_e & \text{if } A \Delta B = e \in E; \\ 0 & \text{otherwise.} \end{cases} \quad ((A,B) \in \left(\binom{V}{k}\right)_2).$$

If we set $f_e^{\{k\}}(A) \equiv \{f_e(a) \,:\, a \in A\}$, we can see that $q(A,B) = w_e$ if $B = f_e^{\{k\}}(A)$ for some (necessarily unique) $e \in E$ and $q(A,B) = 0$ otherwise. The generator of this process is very similar to the one for $\mathsf{RW}(G)$:

$$Q_{\mathsf{EX}(k,G)}(p) = \sum_{e \in E} w_e \, (p \circ f_e^{\{k\}} - p) \ (p \in \mathbb{R}^{\binom{V}{k}}). \tag{3.1.2}$$

Moreover, since the transition rates are symmetric, the stationary distribution of $\mathsf{EX}(k,G)$ is $\mathrm{Unif}\left(\binom{V}{k}\right)$.

*The* $k$*-particle interchange process on* $G$, denoted by $\mathsf{IP}(k,G)$, has state space $(V)_k$. Define, for $\mathbf{x} = (\mathbf{x}(1), \ldots, \mathbf{x}(k)) \in (V)_k$, $f_e^{(k)}(\mathbf{x}) = (f_e(\mathbf{x}(1)), \ldots, f_e(\mathbf{x}(k)))$. The transition rates of $\mathsf{IP}(k,G)$ are given by $q^{(k)}(\mathbf{x},\mathbf{y}) = w_e$ if $\mathbf{y} = f_e(\mathbf{x})$ and 0 otherwise (for $(\mathbf{x},\mathbf{y}) \in ((V)_k)_2$. The generator of this process is:

$$Q_{\mathsf{IP}(k,G)}(g) = \sum_{e \in E} w_e \, (g \circ f_e^{(k)} - g) \ (g \in \mathbb{R}^{(V)_k}). \tag{3.1.3}$$

This process also has symmetric transition rates, and its stationary distribution is $\mathrm{Unif}\left((V)_k\right)$.

**Notational convention 3.1.1** *For simplicitly, the maps $f_e, f_e^{\{k\}}$ and $f_e^{(k)}$ will all be denoted by $f_e$ in what follows.*

## 3.2 The standard graphical construction

We now present the standard *graphical construction* of these three processes. Graphical constructions are standard tools in the study of Interacting Particle Systems [15] and are usually attributed to Harris in the literature. The basic construction presented here will be elaborated upon later in the paper; see Section 5.

Set $W = \sum_{e \in E} w_e$. We need a marked Poisson process, ie. a pair of independent ingredients given as follows:

1. A Poisson process $\mathcal{P} = \{\tau_1 \leq \tau_2 \leq \tau_3 \leq \dots\} \subset [0, +\infty)$ with rate $W$.
2. An i.i.d. sequence of $E$-valued random variables ("markings") $\{e_n\}_{n \in \mathbb{N}}$, with

$$\forall n \in \mathbb{N},\ \mathbb{P}\left(e_n = e\right) = w_e / W.$$

Let $0 \leq t \leq s < +\infty$ be given. We define a random permutation $I_{(t,s]} : V \to V$ associated with the time interval $(t, s]$ as follows: if $\mathcal{P} \cap (t, s] = \emptyset$, $I_{(t,s]}$ is the identity map on $V$. If on the other hand

$$\mathcal{P} \cap (t, s] \equiv \{\tau_j \,:\, m \leq j \leq n\} \neq \emptyset,$$

we set $I_{(t,s]} = f_{e_n} \circ f_{e_{n-1}} \circ \cdots \circ f_{e_m}$; ie. $I_{(t,s]}$ is the composition of each transposition $f_{e_j}$ corresponding to $\tau_j \in (t, s]$, and the transpositions are composed in the order they appear. We also set $I_t \equiv I_{(0,t]}$ for $t > 0$ and $I_{(t,t]}$ =identity map over $V$.

**Remark 3.2.1** *Strictly speaking, we should worry about what happens if $\mathcal{P} \cap (t, s]$ is infinite, or (more generally) some finite interval $(a, b]$ in $[0, +\infty)$ has infinite intersection with $\mathcal{P}$. However, since the probability of any of this holding is 0, we will simply ignore these issues.*

Notice the following simple properties:

**Proposition 3.2.1 (Proof omitted)** *For all $0 \leq t \leq s \leq r$, $I_{(t,r]} = I_{(s,r]} \circ I_{(t,s]}$.*

**Proposition 3.2.2** *For all $0 \leq t \leq s < +\infty$, $I_{(t,s]}$ and $I_{(t,s]}^{-1}$ have the same distribution.*

*Proof:* [Sketch] $\mathcal{S} \equiv \mathcal{P} \cap (t, s]$ can be generated as follows: sample $\{\theta_i\}_{i \geq 1}$ i.i.d. uniform from $(t, s]$ and set $\mathcal{S} \equiv \{\theta_i\}_{i \leq N}$, where $N$ is a r.v. independent from the $\theta_i$ whose distribution is Poisson with mean $s - t$. For each $i$, $\theta_i$ and $s - \theta_i + t$ have the same distribution, hence the random set

$$\mathcal{S}' \equiv \{s - \theta_i + t \; 1 \leq i \leq N\} \text{ (with } N \text{ as before)}$$

has the same law as $\mathcal{S}$. Moreover, the points of $\mathcal{S}$ appear in $\mathcal{S}'$ in reversed order.

$I^{-1}_{(t,s]}$ can be obtained by reversing the order of the compositions in the construction of $I_{(t,s]}$. This reversal corresponds to replacing $\mathcal{S}$ with $\mathcal{S}'$ – which does not change the distribution, as just shown – and also reversing the order of the markings – which also does not change the distribution. It follows that $I^{-1}_{(t,s]}$ and $I_{(t,s]}$ have the same law. $\square$

**Proposition 3.2.3** *Let $0 \leq t_0 < t_1 < t_2 < \cdots < t_k$. Then the maps $I_{(t_{i-1},t_i]}$, $1 \leq i \leq k$, are independent.*

*Proof:* [Sketch] Due to standard properties of the Poisson process, the random sets

$$\{\mathcal{P} \cap (t_{i-1}, t_i]\}_{i=1}^k$$

are independent, and so are the markings associated with each of these. $\square$

**Notational convention 3.2.1** *Like in Notational convention 3.1.1, we "lift" the random maps $I_{(t,s]}$ to permutations of $\binom{V}{k}$ and $(V)_k$, which we also denote by $I_{(t,s]}$:*

$$\begin{aligned} I_{(t,s]}(A) &\equiv \{I_{(t,s]}(a) \,:\, a \in A\} \; \left(A \in \binom{V}{k}\right); \\ I_{(t,s]}(\mathbf{x}) &\equiv (I_{(t,s]}(\mathbf{x}(1)), I_{(t,s]}(\mathbf{x}(2)), \ldots, I_{(t,s]}(\mathbf{x}(k))) \;\; (\mathbf{x} \in (V)_k)\,. \end{aligned}$$

*For brevity, we will often write $x^I_t, A^I_t, \mathbf{x}^I_t$ instead of $I_t(x), I_t(A), I_t(\mathbf{x})$ (resp.).*

The key property of the graphical construction is:

**Proposition 3.2.4** *Let $t_0 \geq 0$. Then:*

1. *For each $x \in V$, the process $\{I_{(t_0,t+t_0]}(x)\}_{t \geq 0}$ is a realization of $\mathsf{RW}(G)$ with initial state $x$.*
2. *For each $A \in \binom{V}{k}$, the process $\{I_{(t_0,t+t_0]}(A)\}_{t \geq 0}$ is a realization of $\mathsf{EX}(k, G)$ with initial state $A$.*
3. *For each $\mathbf{x} \in (V)_k$, the process $\{I_{(t_0,t+t_0]}(\mathbf{x})\}_{t \geq 0}$ is a realization of $\mathsf{IP}(k, G)$ with initial state $\mathbf{x}$.*

*Proof:* [Sketch] This is a standard result in the Interacting Particle Systems literature, so we only sketch the proof of the third statement when $t_0 = 0$. Suppose we are given the whole trajectory $\{I_s(\mathbf{x})\}_{0\le s\le t}$, and that moreover $I_t(\mathbf{x}) = \mathbf{y}$. By Proposition 3.2.1, the conditional probability that $I_{t+\epsilon}(\mathbf{x}) = f_e(\mathbf{y})$ (where $f_e(\mathbf{y}) \neq \mathbf{y}$) is the conditional probability of $I_{(t,t+\epsilon]}(\mathbf{y}) = f_e(\mathbf{y})$. This is in fact an unconditional probability, since $I_{(t,t+\epsilon]}$ is independent of $\{I_s(\mathbf{x})\}_{0\le s\le t}$ (easy consequence of Proposition 3.2.3). Moreover, for $\epsilon > 0$ small, the probability of $I_{(t,t+\epsilon]}(\mathbf{y}) = f_e(\mathbf{y})$ equals

$$\mathbb{P}\left(I_{(t,t+\epsilon]} = f_e\right) = \mathbb{P}\left(\exists n \in \mathbb{N} : \mathcal{P} \cap (t, t+\epsilon] = \{\tau_n\} \text{ and } e_n = w_e\right) + O\left(\epsilon^2\right),$$

since the probability of $\mathcal{P} \cap (t, t+\epsilon]$ having two or more points is $O\left(\epsilon^2\right)$. Since the marking $e_n$ is independent of $\mathcal{P}$ and $|\mathcal{P} \cap (t, t+\epsilon]|$ is a Poisson r.v. with mean $\epsilon W$, we have:

$$\begin{aligned}&\mathbb{P}\left(\exists n \in \mathbb{N} : \mathcal{P} \cap (t, t+\epsilon] = \{\tau_n\} \text{ and } e_n = w_e\right)\\ &\qquad = \mathbb{P}\left(\exists n \in \mathbb{N} : \mathcal{P} \cap (t, t+\epsilon] = \{\tau_n\}\right) \frac{w_e}{W} = (\epsilon w_e)\, e^{-W\epsilon} = \epsilon w_e + O\left(\epsilon^2\right).\end{aligned}$$

We deduce that $\mathbb{P}\left(I_{(t,t+\epsilon]} = f_e\right) = \epsilon w_e + O\left(\epsilon^2\right)$. That is, the probability of a transition from $I_t(\mathbf{x}) = \mathbf{y}$ to $f_e(\mathbf{y})$ given the whole past is $\epsilon w_e$, precisely what one should get in $\mathsf{IP}(k, G)$. □

## 3.3 The negative correlation property

$\mathsf{EX}(k, G)$ enjoys important *negative correlation properties*. In this paper we only need a very special result, which is contained in any of [15, 14, 3].

**Lemma 3.3.1** *Given $A \in \binom{V}{k}$, let $\{A_t^I\}_{t\ge 0}$ be a realization of $\mathsf{EX}(k, G)$ starting from $A$. Then for all $\mathbf{u} \in (V)_2$ – ie. for all distinct $\mathbf{u}(1), \mathbf{u}(2) \in V$ –, we have:*

$$\mathbb{P}\left(\{\mathbf{u}(1) \in A_t^I\} \cap \{\mathbf{u}(2) \in A_t^I\}\right) \le \mathbb{P}\left(\mathbf{u}(1) \in A_t^I\right) \mathbb{P}\left(\mathbf{u}(2) \in A_t^I\right).$$

Let us rewrite this result in another convenient form, using the construction in the previous section.

**Corollary 3.3.1** *Given $\mathbf{u} \in (V)_2$, let $\{\mathbf{u}_t^I\}_{t\ge 0}$ be a realization of $\mathsf{IP}(2, G)$ starting from $\mathbf{u}$. Then for all $A \subset V$:*

$$\mathbb{P}\left(\{\mathbf{u}_t^I(1) \in A\} \cap \{\mathbf{u}_t^I(2) \in A\}\right) \le \mathbb{P}\left(\mathbf{u}_t^I(1) \in A\right) \mathbb{P}\left(\mathbf{u}_t^I(2) \in A\right).$$

*Proof:* We use Proposition 3.2.4 and write $A_t^I \equiv I_t(A)$. The inequality in Lemma 3.3.1 can be rewritten as:

$$\mathbb{P}\left(\{I_t^{-1}(\mathbf{u}(1)) \in A\} \cap \{I_t^{-1}(\mathbf{u}(2)) \in A\}\right) \le \mathbb{P}\left(I_t^{-1}(\mathbf{u}(1)) \in A\right) \mathbb{P}\left(I_t^{-1}(\mathbf{u}(2)) \in A\right).$$

$I_t^{-1}$ and $I_t$ have the same distribution (Proposition 3.2.2), hence:

$$\mathbb{P}\left(\{I_t(\mathbf{u}(1)) \in A\} \cap \{I_t(\mathbf{u}(2)) \in A\}\right) \leq \mathbb{P}\left(I_t(\mathbf{u}(1)) \in A\right) \mathbb{P}\left(I_t(\mathbf{u}(2)) \in A\right).$$

Since $\mathbf{u}_t^I \equiv (\mathbf{u}_t^I(1), \mathbf{u}_t^I(2)) = (I_t(\mathbf{u}(1)), I_t(\mathbf{u}(2)))$, this finishes the proof. □

# 4 The dynamics of pairs of particles

The goal of this section is to prove Lemma 1.4.1. We fix a weighted graph $G = (V, E, \{w_e\}_{e \in E})$ for the remainder of the section (and of the paper). The definitions of $\mathsf{RW}(G)$, $\mathsf{RW}(k, G)$, $\mathsf{EX}(k, G)$ and $\mathsf{IP}(k, G)$ are all relative to this graph.

## 4.1 Some facts on $\mathsf{RW}(2, G)$ and $\mathsf{IP}(2, G)$

Much of this section will involve comparisons between $\mathsf{IP}(2, G)$ and $\mathsf{RW}(2, G)$. The following notational convention will be useful.

**Notational convention 4.1.1** *Given* $\mathbf{x} \in V^2$,

$$\{\mathbf{x}_t^R \equiv (\mathbf{x}_t^R(1), \mathbf{x}_t^R(2)) \,:\, t \geq 0\}$$

*denotes a realization of* $\mathsf{RW}(2, G)$ *from initial state* $\mathbf{x}$. *That is, the trajectories of* $\mathbf{x}_t^R(1), \mathbf{x}_t^R(2)$ *are independent realizations of* $\mathsf{RW}(G)$ *with respective initial states* $\mathbf{x}(1), \mathbf{x}(2)$.

We collect several simple facts about $\mathsf{RW}(2, G)$ and $\mathsf{IP}(2, G)$ that we will need later on.

**Proposition 4.1.1** *The marginal laws of* $\mathbf{x}_t^R(i), \mathbf{x}_t^I(i)$ *are always the same.*

*Proof:* This follows trivially from the graphical construction (cf. Proposition 3.2.4). □

**Proposition 4.1.2** *The mixing times of* $\mathsf{RW}(2, G)$ *satisfy:*

$$\forall \epsilon > 0,\ \mathsf{T}_{\mathsf{RW}(2,G)}(\epsilon) \leq \mathsf{T}_{\mathsf{RW}(G)}(\epsilon/2).$$

*Proof:* Let $\mathrm{Unif}(V)$ and $\mathrm{Unif}(V^2)$ denote the uniform distributions over $V$ and $V^2$ (resp.) Also set $t = \mathsf{T}_{\mathsf{RW}(G)}(\epsilon/2)$. For any $\mathbf{x} \in (V)_2$ we have $d_{\mathrm{TV}}(\mathbf{x}_t^R(i), \mathrm{Unif}(V)) \leq \epsilon/2$ by the choice of $t$. Equation (2.2.5) implies:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^R], \mathrm{Unif}(V^2)) = d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^R(1)] \times \mathcal{L}[\mathbf{x}_t^R(2)], \mathrm{Unif}(V) \times \mathrm{Unif}(V)) \leq \epsilon.$$

Since $\mathbf{x}$ is arbitrary, this finishes the proof. □

**Proposition 4.1.3** *Let $k \in \mathbb{N}$ be given. Then*

$$\forall \epsilon > 0,\ \mathsf{T}_{\mathsf{RW}(2,G)}(2^{-k}) \leq (k+1)\mathsf{T}_{\mathsf{RW}(G)}(1/4).$$

*Proof:* Follows from the $\mathsf{T}_{\mathsf{RW}(2,G)}(2^{-k}) \leq \mathsf{T}_{\mathsf{RW}(G)}(2^{-k-1})$ (previous Proposition) and from

$$\mathsf{T}_{\mathsf{RW}(G)}(2^{-k-1}) \leq (k+1)\,\mathsf{T}_{\mathsf{RW}(G)}(1/4),$$

a direct consequence of Proposition 2.3.1. □

The next Lemma has the following meaning. Suppose $t$ is so large that $\mathbf{x}_t^R$ is close to equilibrium. In this case, $\mathbb{E}\left[\phi(\mathbf{x}_t^R)\right]$ is close to the uniform average of $\phi$ over $V^2$, for all mappings $0 \leq \phi \leq 1$. The Lemma shows that if $\phi$ is *symmetric*, $\mathbb{E}\left[\phi(\mathbf{x}_t^I)\right]$ cannot be much larger than that average. The negative correlation property proven in Corollary 3.3.1 is invoked at a key step of the proof of the Lemma.

**Lemma 4.1.1** *Let $\phi : V^2 \to [0,1]$ be symmetric, in the sense that*

$$\forall \mathbf{u} \in (V)_2,\ \phi(\mathbf{u}(1), \mathbf{u}(2)) = \phi(\mathbf{u}(2), \mathbf{u}(1)).$$

*Then:*

$$\forall \epsilon \in (0, 1/64),\ \forall t \geq \mathsf{T}_{\mathsf{RW}(G)}(\epsilon),\ \forall \mathbf{x} \in (V)_2,\ \mathbb{E}\left[\phi(\mathbf{x}_t^I)\right] \leq 8\sqrt{\epsilon} + 9 \sum_{\mathbf{v} \in V^2} \frac{\phi(\mathbf{v})}{|V|^2}.$$

*Proof:* Define

$$G \equiv \left\{ a \in V \,:\, \max_{i=1,2} \left| \mathbb{P}\left(\mathbf{x}_t^I(i) = a\right) - \frac{1}{|V|} \right| \leq \frac{2\sqrt{\epsilon}}{|V|} \right\}$$

and $B = V \backslash G$. We will show towards the end of the proof that:

$$\mathbb{P}\left(\mathbf{x}_t^I \notin G^2\right) \leq 8\sqrt{\epsilon}, \tag{4.1.1}$$

which (since $0 \leq \phi \leq 1$) implies:

$$\mathbb{E}\left[\phi(\mathbf{x}_t^I)\mathbb{I}_{(V)_2 \backslash G^2}(\mathbf{x}_t^I)\right] \leq 8\sqrt{\epsilon}, \tag{4.1.2}$$

On the other hand, notice that:

$$\begin{aligned}
\mathbb{E}\left[\phi(\mathbf{x}_t^I)\mathbb{I}_{G^2}(\mathbf{x}_t^I)\right] &= \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \mathbb{P}\left(\mathbf{x}_t^I = (\mathbf{a}(1),\mathbf{a}(2))\right)\ \phi(\mathbf{a}) \\
\text{(symmetry)} &= \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \left(\frac{\mathbb{P}\left(\mathbf{x}_t^I = (\mathbf{a}(1),\mathbf{a}(2))\right) + \mathbb{P}\left(\mathbf{x}_t^I = (\mathbf{a}(2),\mathbf{a}(1))\right)}{2}\right)\ \phi(\mathbf{a}) \\
&= \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \mathbb{P}\left(\mathbf{x}_t^I(1)\in\{\mathbf{a}(1),\mathbf{a}(2)\},\ \mathbf{x}_t^I(2)\in\{\mathbf{a}(1),\mathbf{a}(2)\}\right)\ \phi(\mathbf{a}) \\
\text{(Cor. 3.3.1)} &\leq \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \mathbb{P}\left(\mathbf{x}_t^I(1)\in\{\mathbf{a}(1),\mathbf{a}(2)\}\right)\ \mathbb{P}\left(\mathbf{x}_t^I(2)\in\{\mathbf{a}(1),\mathbf{a}(2)\}\right)\ \phi(\mathbf{a}) \\
\text{(Prop. 4.1.1)} &= \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \mathbb{P}\left(\mathbf{x}_t^R(1)\in\{\mathbf{a}(1),\mathbf{a}(2)\}\right)\ \mathbb{P}\left(\mathbf{x}_t^R(2)\in\{\mathbf{a}(1),\mathbf{a}(2)\}\right)\ \phi(\mathbf{a}) \\
(\mathbf{a}(i)\in G) &\leq \sum_{\mathbf{a}\in(G^2)\cap(V)_2} \left(\frac{2+2\sqrt{\epsilon}}{|V|}\right)^2\ \phi(\mathbf{a}) \\
(\epsilon\leq 1/4) &\leq 9\sum_{\mathbf{a}\in V^2}\frac{\phi(\mathbf{a})}{|V|^2}.
\end{aligned}$$

Combining this with (4.1.2) finishes the proof, except for (4.1.1). To prove that, we first notice that:

$$\frac{\sqrt{\epsilon}|B|}{|V|} \leq \frac{1}{2}\sum_{a\in V}\left|\mathbb{P}\left(\mathbf{x}_t^R(1)=a\right)-\frac{1}{|V|}\right| + \left|\mathbb{P}\left(\mathbf{x}_t^R(2)=a\right)-\frac{1}{|V|}\right|$$

as each $a\in B$ contributes at least $\sqrt{\epsilon}/|V|$ to the sum. But the RHS equals:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^R(1)],\mathrm{Unif}(V)) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^R(2)],\mathrm{Unif}(V)) \leq 2\epsilon$$

since $t\geq \mathsf{T}_{\mathsf{RW}(G)}(\epsilon)$. We deduce:

$$\frac{\sqrt{\epsilon}|B|}{|V|}\leq 2\epsilon,\ \text{or equivalently } |G|\geq(1-2\sqrt{\epsilon})\,|V|.$$

Moreover, $\mathbb{P}\left(\mathbf{x}_t^R(i)=a\right)\geq(1-2\sqrt{\epsilon})|V|^{-1}$ for all $a\in G$, hence:

$$\mathbb{P}\left(\mathbf{x}_t^R(i)\in G\right)\geq\frac{|G|}{|V|}\,(1-2\sqrt{\epsilon})\geq(1-2\sqrt{\epsilon})^2\geq 1-4\sqrt{\epsilon}.$$

We conclude that:

$$\begin{aligned}
\mathbb{P}\left(\mathbf{x}_t^I\notin G^2\right) &\leq \mathbb{P}\left(\mathbf{x}_t^I(1)\notin G\right)+\mathbb{P}\left(\mathbf{x}_t^I(2)\notin G\right) \\
\text{(Proposition 4.1.1)} &= \mathbb{P}\left(\mathbf{x}_t^R(1)\notin G\right)+\mathbb{P}\left(\mathbf{x}_t^R(2)\notin G\right) \\
\text{(previous ineq.)} &\leq 8\sqrt{\epsilon}.
\end{aligned}$$

Thus (4.1.1) holds and we are done. □

## 4.2 When collisions are nearly as fast as mixing

Recalling Notational convention 4.1.1, we define the first *meeting time* $M(\mathbf{x})$ of $\mathsf{RW}(2,G)$ started from $\mathbf{x} \in V^2$ as the smallest $t_0 \geq 0$ such that $\mathbf{x}^R_{t_0}(1) = \mathbf{x}^R_{t_0}(2)$ (this is a.s. finite by ergodicity). We will also write

$$M_{\geq t}(\mathbf{x}) = \inf\{h_0 \geq 0 \,:\, \mathbf{x}^R_{t+h_0}(1) = \mathbf{x}^R_{t+h_0}(2)\}$$

for the time until the first meeting after $t$.

The following definition will be crucial for our analysis.

**Definition 1** *We say that a weighted graph $G$ is easy if:*

$$\sup_{\mathbf{x}\in V^2} \mathbb{P}\left(M(\mathbf{x}) > 20000\,\mathsf{T}_{\mathsf{RW}(G)}(1/4)\right) \leq 1/8.$$

Notice that this definition is not vacuous, as all long enough paths and cycles are easy. The next Proposition proves Lemma 1.4.1 for all easy graphs via a coupling argument due to Aldous and Fill.

**Proposition 4.2.1** *Lemma 1.4.1 holds for all easy weighted graphs.*

*Proof:* [Sketch] Given $G$, Aldous and Fill [2, Chapter 14, Section 5] construct a coupling of $\mathsf{IP}(|V|,G)$ started from two different states $\mathbf{u},\mathbf{v}$. Letting $\{\mathbf{u}^C_t, \mathbf{v}^C_t\}_{t\geq 0}$ denote the coupled trajectories, the following property holds: for each $1 \leq i \leq |V|$, $\mathbf{u}^C_t(i), \mathbf{v}^C_t(i)$ behave as independent random walks up to their first time $M_i$, and then move together. This implies:

$$\forall t \geq 0,\, d_{\mathrm{TV}}(\mathcal{L}[\mathbf{u}^I_t], \mathcal{L}[\mathbf{v}^I_t]) \leq \mathbb{P}\left(\mathbf{u}^C_t \neq \mathbf{v}^C_t\right) \leq \sum_{i=1}^{|V|} \mathbb{P}\left(\mathbf{u}^C_t(i) \neq \mathbf{v}^C_t(i)\right) \leq \sum_{i=1}^{|V|} \mathbb{P}\left(M_i > t\right).$$

It is easy to adapt this to a coupling of $\mathsf{IP}(2,G)$ starting from given $\mathbf{x},\mathbf{y} \in (V)_2$, so that, if $\{\mathbf{x}^C_t, \mathbf{y}^C_t\}_{t\geq 0}$ denotes the coupled trajectories, we have:

$$\forall t \geq 0,\, d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_t], \mathcal{L}[\mathbf{y}^I_t]) \leq \mathbb{P}\left(M_1 > t\right) + \mathbb{P}\left(M_2 > t\right),$$

where $M_i$ is the first meeting time of $\mathbf{x}^C(i)$ and $\mathbf{y}^C(i)$, $i = 1,2$. Now both $M_1$ and $M_2$ are the meeting times of independent random walkers on $G$, which shows that:

$$\forall t \geq 0 \quad \sup_{\mathbf{x},\mathbf{y}\in(V)_2} d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_t], \mathcal{L}[\mathbf{y}^I_t]) \leq 2 \sup_{\mathbf{z}\in V^2} \mathbb{P}\left(M(\mathbf{z}) > t\right).$$

For $t = 20000\,\mathsf{T}_{\mathsf{RW}(G)}(1/4)$ and $G$ easy, the RHS is $\leq 1/4$. By convexity, this implies that:

$$\sup_{\mathbf{x}\in(V)_2} d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_t], \mathrm{Unif}((V)_2)) \leq \frac{1}{4}.$$

In other words, $\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \leq 20000\,\mathsf{T}_{\mathsf{RW}(G)}(1/4)$. □

**Remark 4.2.1** *The argument of Aldous and Fill actually proves Theorem 1.1.1 for all easy graphs; see [2, Chapter 14, Section 5] for details.*

## 4.3 Long time to meet in non-easy graphs

We now consider what happens when $\mathsf{IP}(2, G)$ is performed on a graph that is not easy. Our first goal is to show that independent random walkers take a relatively long time to meet from *most* initial states in $V$.

**Proposition 4.3.1** *Assume $G = (V, E, \{w_e\}_{e \in E})$ is not easy. Then:*

$$\frac{1}{|V|^2} \sum_{\mathbf{v} \in V^2} \mathbb{P}\left(M(\mathbf{v}) \leq 20 \mathsf{T}_{\mathsf{RW}(G)}(1/4)\right) \leq \frac{1}{125}.$$

**Remark 4.3.1** *In general we cannot guarantee that $\mathbb{P}\left(M(\mathbf{v}) < 20 \mathsf{T}_{\mathsf{RW}(G)}(1/4)\right)$ is uniformly small over all $\mathbf{v} \in (V)_2$. For example, a large 3-regular expander is not easy and has $\mathsf{T}_{\mathsf{RW}(G)}(1/4) = \Theta(\ln |V|)$, but if $\mathbf{v} = (\mathbf{v}(1), \mathbf{v}(2))$ with $\mathbf{v}(1)$ and $\mathbf{v}(2)$ adjacent, $\mathbb{P}(M(\mathbf{v}) \leq 1) \geq (1 - e^{-6})/3$.*

*Proof:* [of the Proposition] Set $T = \mathsf{T}_{\mathsf{RW}(G)}(1/4)$. Since $G$ is not easy, there exists some $\mathbf{x} \in V^2$ with:

$$\mathbb{P}(M(\mathbf{x}) > 20000\, T) > 1/8. \tag{4.3.1}$$

Consider some $k \in \mathbb{N}$. Using the Markov property and the notation introduced in Section 4.2, one can write:

$$\mathbb{P}(M(\mathbf{x}) > 40kT) = \mathbb{E}\left[\mathbb{I}_{\{M(\mathbf{x}) > 40(k-1)T\}} \mathbb{P}\left(M_{\geq 40(k-1)T}(\mathbf{x}) > 40T \mid \mathbf{x}^R_{40(k-1)T}\right)\right].$$

The Markov property implies:

$$\begin{aligned}
&\mathbb{P}\left(M_{\geq 40(k-1)T}(\mathbf{x}) > 40T \mid \mathbf{x}^R_{40(k-1)T} = \mathbf{y}\right) \\
&\qquad = \mathbb{P}\left(\forall 0 \leq t \leq 40T,\, \mathbf{y}^R_t(1) \neq \mathbf{y}^R_t(2)\right) = \mathbb{P}(M(\mathbf{y}) > 40T),
\end{aligned}$$

hence:

$$\begin{aligned}
\mathbb{P}(M(\mathbf{x}) > 40kT) &\leq \left(\sup_{\mathbf{y} \in V^2} \mathbb{P}(M(\mathbf{y}) > 40T)\right) \mathbb{E}\left[\mathbb{I}_{\{M(\mathbf{x}) > 40(k-1)T\}}\right] \\
&\leq \left(\sup_{\mathbf{y} \in V^2} \mathbb{P}(M(\mathbf{y}) > 40T)\right) \mathbb{P}(M(\mathbf{x}) > 40(k-1)T) \\
(\text{...induction...}) &\leq \left(\sup_{\mathbf{y} \in V^2} \mathbb{P}(M(\mathbf{y}) > 40T)\right)^k.
\end{aligned}$$

Applying this to $k = 500$ and using the bound in (4.3.1) gives the following with room to spare:

$$\sup_{\mathbf{y}\in V^2} \mathbb{P}\left(M(\mathbf{y}) > 40T\right) \geq 8^{-1/500} \geq e^{-3/500} \geq \frac{497}{500}.$$

Fix some $\mathbf{y} \in V^2$ achieving this supremum. Notice that $M(\mathbf{y}) > 40T$ holds if and only if $\mathbf{y}_t^R(1) \neq \mathbf{y}_t^R(2)$ for all $0 \leq t \leq 40T$. If that is the case, $\mathbf{y}_{20T+h}^R(1) \neq \mathbf{y}_{20T+h}^R(2)$ for all $0 \leq h \leq 20T$. Using the Markov property as before, we see that:

$$\frac{497}{500} \leq \mathbb{P}\left(M(\mathbf{y}) > 40T\right) \leq \mathbb{P}\left(M_{\geq 20T}(\mathbf{y}) > 20T\right) = \sum_{\mathbf{v}\in V^2} \mathbb{P}\left(\mathbf{y}_{20T}^R = \mathbf{v}\right) \mathbb{P}\left(M(\mathbf{v}) > 20T\right).$$

Moreover, by (2.2.2),

$$\sum_{\mathbf{v}\in V^2} \mathbb{P}\left(\mathbf{y}_{20T}^R = \mathbf{v}\right) \mathbb{P}\left(M(\mathbf{v}) > 20T\right) \leq \sum_{\mathbf{v}\in V^2} \frac{\mathbb{P}\left(M(\mathbf{v}) > 20T\right)}{|V|^2} + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_{20T}^R], \mathrm{Unif}(V^2)).$$

Hence:

$$\frac{497}{500} - d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_{20T}^R], \mathrm{Unif}(V^2)) \leq \sum_{\mathbf{v}\in V^2} \frac{\mathbb{P}\left(M(\mathbf{v}) > 20T\right)}{|V|^2}.$$

We finish by noting that, by Proposition 4.1.3, $20T \geq \mathsf{T}_{\mathsf{RW}(2,G)}(2^{-19})$, hence:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_{20T}^R], \mathrm{Unif}(V^2)) \leq 2^{-19} \leq \frac{1}{500},$$

and therefore

$$\sum_{\mathbf{v}\in V^2} \frac{\mathbb{P}\left(M(\mathbf{v}) > 20T\right)}{|V|^2} \geq \frac{496}{500} = 1 - \frac{1}{125}.$$

□

## 4.4 If meeting takes a long time, $\mathsf{IP}(2, G)$ and $\mathsf{RW}(2, G)$ are similar

Proposition 4.3.1 shows that from most initial states, the meeting time is unlikely to be smaller than $20\mathsf{T}_{\mathsf{RW}(G)}(1/4)$. We now show that $\mathsf{IP}(2, G)$ is similar to $\mathsf{RW}(2, G)$ essentially until the first meeting time.

**Proposition 4.4.1** *For any* $\mathbf{x} \in (V)_2$ *and* $s \geq 0$,

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_s^R], \mathcal{L}[\mathbf{x}_s^I]) \leq \mathbb{P}\left(M(\mathbf{x}) \leq s\right).$$

We will only need the following simple corollary (proof omitted) in what follows.

**Corollary 4.4.1** *For any* $\mathbf{x}, \mathbf{y} \in (V)_2$ *and* $s \geq 0$,

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_s^I], \mathcal{L}[\mathbf{y}_s^I]) \leq \mathbb{P}\left(M(\mathbf{x}) \leq s\right) + \mathbb{P}\left(M(\mathbf{y}) \leq s\right) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_s^R], \mathcal{L}[\mathbf{y}_s^R]).$$

*Proof:* [of Proposition 4.4.1] We present a coupling of $\{\mathbf{x}_t^I\}_{t\geq 0}$ and $\{\mathbf{x}_t^R\}_{t\geq 0}$ such that the two processes agree up to $M(\mathbf{x})$. The Proposition then follows from the coupling characterization of $d_{\mathrm{TV}}(\cdot, \cdot\cdot)$, cf. Section 2.2.

Our coupling is given by a continuous-times Markov chain on $S = V^2 \times (V)_2$ with transition rates given by $q(\cdot, \cdot\cdot)$. The state space can be split into two parts, $\Delta \equiv \{(\mathbf{z}, \mathbf{z}) : \mathbf{z} \in (V)_2\}$ and its complement $\Delta^c$.

- **Transition rule 1:** The transition rates from any pair $(\mathbf{x}, \mathbf{y}) \in \Delta^c$ to any other pair in $S$ are the same as those of independent realizations of $\mathsf{RW}(2, G)$ and $\mathsf{IP}(2, G)$.
- **Transition rule 2:** The transition rates from a pair $(\mathbf{x}, \mathbf{x}) \in \Delta$ are determined as follows:
  - **Transition rule 2.1:** For each $e \in E$ with $|e \cap \{\mathbf{x}(1), \mathbf{x}(2)\}| = 1$,
  $$q((\mathbf{x}, \mathbf{x}), (f_e(\mathbf{x}), f_e(\mathbf{x}))) = w_e;$$
  - **Transition rule 2.2:** If $e \in E$ satisfies $e = \{\mathbf{x}(1), \mathbf{x}(2)\}$,
  $$\begin{cases} q((\mathbf{x}, \mathbf{x}), (f_e(\mathbf{x}), (\mathbf{x}(1), \mathbf{x}(1)))) = w_e, \\ q((\mathbf{x}, \mathbf{x}), (\mathbf{x}, (\mathbf{x}(2), \mathbf{x}(2)))) = w_e; \end{cases}$$
  - **Transition rule 2.3:** All other potential transitions have rate 0.

It should be clear that the first marginals of the transition rates give $\mathsf{IP}(2, G)$. Indeed, this is obvious outside of $\Delta$, and inside $\Delta$ one can check for any distinct $\mathbf{x}, \mathbf{y} \in (V)_2$, the marginal transition rate is $w_e$ if $\mathbf{y} = f_e(\mathbf{x})$ for some (necessarily unique) $e \in E$ and 0 otherwise.

We *claim* that the second marginals of the process correspond to $\mathsf{RW}(2, G)$. This is obvious in the set $\Delta^c$. In $\Delta$, consider some $\mathbf{x} \in (V)_2$ and $e \in E$. We have three cases:

- If $e \cap \{\mathbf{x}(1), \mathbf{x}(2)\}$ has one element – say, $\mathbf{x}(1)$ –, $f_e(\mathbf{x}(2)) = \mathbf{x}(2)$ and therefore $f_e(\mathbf{x}) = (f_e(\mathbf{x}(1)), \mathbf{x}(2))$. Hence in this case the second marginal of the transition rates are equal to $w_e$ for all transitions of the form $\mathbf{x} \to (f_e(\mathbf{x}(1)), \mathbf{x}(2))$ or (analogously) $\mathbf{x} \to (\mathbf{x}(1), f_e(\mathbf{x}(2)))$. This agrees with $\mathsf{RW}(2, G)$.
- If $e = \{\mathbf{x}(1), \mathbf{x}(2)\}$, the second marginal of the transition rates equals $w_e$ for both possible transitions:
  $$\mathbf{x} \to (\mathbf{x}(1), \mathbf{x}(1)) = (\mathbf{x}(1), f_e(\mathbf{x}(1))) \text{ and } \mathbf{x} \to (\mathbf{x}(2), \mathbf{x}(2)) = (f_e(\mathbf{x}(1)), \mathbf{x}(2)).$$
  This also agrees with $\mathsf{RW}(2, G)$.

- Finally, all other potential transitions have marginal rate 0 for $q$ and rate 0 for $\mathsf{RW}(2, G)$.

We have thus shown that an evolution of the Markov Chain with transition rates $q(\cdot, \cdot\cdot)$ as above, started from a state $(\mathbf{x}, \mathbf{x}) \in \Delta$, is a coupling of $(\mathbf{x}_t^I)_{t\geq 0}$ and $(\mathbf{x}_t^R)_{t\geq 0}$. The first time at which $\mathbf{x}_t^I \neq \mathbf{x}_t^R$ is the first exit time for $\Delta$. The transition rules for $q$ show that this will happen at the first time a transition occurs according to rule 2.2 above. But this is precisely the first time at which $\mathbf{x}_t^R(1) = \mathbf{x}_t^R(2)$, which is $M(\mathbf{x})$. This finishes the proof. □

## 4.5 Proof of the mixing time bound for $\mathsf{IP}(2, G)$

We now use the tools developed above in order to prove Lemma 1.4.1.
*Proof:* [of Lemma 1.4.1] The case of easy graphs is covered by Proposition 4.2.1, so assume $G = (V, E, \{w_e\}_{e\in E})$ is not easy. Let $\mathbf{x}, \mathbf{y}$ be given and $T \equiv \mathsf{T}_{\mathsf{RW}(G)}(1/4)$. Notice that for all $A \subset (V)_2$, if $\{\mathbf{x}_t^I\}_{t\geq 0}$, $\{\mathbf{y}_t^I\}_{t\geq 0}$ are defined over the same probability space,

$$\begin{aligned}\mathbb{P}\left(\mathbf{x}_{40T}^I \in A\right) - \mathbb{P}\left(\mathbf{y}_{40T}^I \in A\right) &= \mathbb{E}\left[\mathbb{P}\left(\mathbf{x}_{40T}^I \in A \mid \mathbf{x}_{20T}^I\right) - \mathbb{P}\left(\mathbf{y}_{40T}^I \in A \mid \mathbf{y}_{20T}^I\right)\right] \\ &\leq \mathbb{E}\left[d_{\mathrm{TV}}(\mathbb{P}\left(\mathbf{x}_{40T}^I \in \cdot \mid \mathbf{x}_{20T}^I\right), \mathbb{P}\left(\mathbf{y}_{40T}^I \in \cdot \mid \mathbf{y}_{20T}^I\right))\right].\end{aligned}$$

Maximizing over $A$ yields:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_{40T}^I], \mathcal{L}[\mathbf{y}_{40T}^I]) \leq \mathbb{E}\left[d_{\mathrm{TV}}(\mathbb{P}\left(\mathbf{x}_{40T}^I \in \cdot \mid \mathbf{x}_{20T}^I\right), \mathbb{P}\left(\mathbf{y}_{40T}^I \in \cdot \mid \mathbf{y}_{20T}^I\right))\right]. \tag{4.5.1}$$

Condition on $\mathbf{x}_{20T}^I = \mathbf{v}, \mathbf{y}_{20T}^I = \mathbf{w}$ for $\mathbf{v}, \mathbf{w} \in (V)_2$. By the Markov property and Corollary 4.4.1,

$$\begin{aligned}d_{\mathrm{TV}}(\mathbb{P}\left(\mathbf{x}_{40T}^I \in \cdot \mid \mathbf{x}_{20T}^I = \mathbf{v}\right), \mathbb{P}\left(\mathbf{y}_{40T}^I \in \cdot \mid \mathbf{y}_{20T}^I = \mathbf{w}\right)) &= d_{\mathrm{TV}}(\mathcal{L}[\mathbf{v}_{20T}^I], \mathcal{L}[\mathbf{w}_{20T}^I]) \\ &\leq \mathbb{P}\left(M(\mathbf{v}) \leq 20T\right) + \\ &\quad + \mathbb{P}\left(M(\mathbf{w}) \leq 20T\right) + \\ &\quad + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{v}_{20T}^R], \mathcal{L}[\mathbf{w}_{20T}^R]).\end{aligned}$$

Proposition 4.1.3 implies the third term in the RHS is $\leq 2^{-19}$ for any $\mathbf{v}, \mathbf{w}$. Using this in conjunction with (4.5.1), we obtain:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_{40T}^I], \mathcal{L}[\mathbf{y}_{40T}^I]) \leq \mathbb{E}\left[\phi(\mathbf{x}_{20T}^I)\right] + \mathbb{E}\left[\phi(\mathbf{y}_{20T}^I)\right] + 2^{-19} \tag{4.5.2}$$

where $\phi(\mathbf{z}) = \mathbb{P}\left(M(\mathbf{z}) \leq 20T\right)$. Notice that $\phi$ nonnegative and symmetric; that is, the distribution of $M$ does not change if one swaps the initial positions of the two random walkers.

We may apply Lemma 4.1.1 and the fact that $20T \geq \mathsf{T}_{\mathsf{RW}(G)}(2^{-20})$ (cf. Proposition 2.3.1) to deduce:

$$\mathbb{E}\left[\phi(\mathbf{x}^I_{20T})\right] \leq 2^{-7} + 9 \sum_{\mathbf{v}\in V^2} \frac{\mathbb{P}\left(M(\mathbf{v}) \leq 20T\right)}{|V|^2}. \tag{4.5.3}$$

Applying the same reasoning to $\phi(\mathbf{y}^I_{20T})$ and plugging the results into (4.5.2), we obtain:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_{40T}], \mathcal{L}[\mathbf{y}^I_{40T}]) \leq 18 \sum_{\mathbf{v}\in V^2} \frac{\mathbb{P}\left(M(\mathbf{v}) \leq 20T\right)}{|V|^2} + 2^{-6} + 2^{-19}. \tag{4.5.4}$$

Finally, we use the fact that $G$ is not easy, combined with Proposition 4.3.1, to deduce:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_{40T}], \mathcal{L}[\mathbf{y}^I_{40T}]) \leq \frac{18}{125} + 2^{-9} + 2^{-6} \leq 1/4 \tag{4.5.5}$$

with room to spare. By convexity,

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_{40T}], \mathrm{Unif}((V)_2)) \leq 1/4, \tag{4.5.6}$$

Since this holds for all $\mathbf{x} \in (V)_2$, we have $\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \leq 40T$, which implies Lemma 1.4.1 for non-easy graphs. □

**Remark 4.5.1** *The first inequality in (4.5.3) follows from Lemma 4.1.1, which is a consequence of the negative correlation property (cf. Lemma 3.3.1 and Corollary 3.3.1). This is the first crucial use we make of negative correlation in this paper.*

# 5 The chameleon process

In the previous section we determined the order of magnitude of the mixing time of $\mathsf{IP}(2, G)$. Going beyond two particles will require an important additional idea that is based on Morris' paper [20]. His idea is to introduce the so-called *chameleon process* to keep track of the conditional distribution of one particle in $\mathsf{IP}(k, G)$. We will need a different process, which will nevertheless call by the same name.

## 5.1 A modified graphical construction

We will need consider a variant of the construction of $\mathsf{IP}(k, G)$ presented in Section 3.2. Consider *three* independent ingredients:

1. A Poisson process $\mathcal{P} = \{\tau_1 \leq \tau_2 \leq \tau_3 \leq \dots\} \subset [0, +\infty)$ with rate $2W$.
2. An i.i.d. sequence of $E$-valued random variables $\{e_n\}_{n\in\mathbb{N}}$, with $\mathbb{P}\left(e_n = e\right) = w_e/W$.

3. An i.i.d. sequence of *coin flips* $\{c_n\}_{n\in\mathbb{N}}$ with $\mathbb{P}(c_n = 1) = \mathbb{P}(c_n = 0) = 1/2$.

Recall the definition of $f_e$ from Section 3.1 (see also Notational convention 3.1.1) and now set $f_e^1 = f_e$, $f_e^0 =$the identity function. We modify the definition of the maps $I_{(t,s]}$ from Section 3.2 as follows: if $\mathcal{P} \cap (t,s] = \emptyset$, $I_{(t,s]}$ is the identity map, as before. Otherwise,

$$\mathcal{P} \cap (t,s] = \{\tau_n < \tau_{n+1} < \cdots < \tau_m\}$$

and we set:

$$I_{(t,s]} = f_{e_m}^{c_m} \circ \cdots \circ f_{e_{n+1}}^{c_{n+1}} \circ f_{e_n}^{c_n}.$$

**Proposition 5.1.1** *The joint distribution of the maps $I_{(t,s]}$, $0 \le t < s < +\infty$, is the same as in Section 3.2.*

*Proof:* Let $\{k_n\}_{n\in\mathbb{N}}$ be the (a.s. infinite) sequence of all $k \in \mathbb{N}$ with $c_k = 1$. The thinning property of the Poisson process implies that:

$$\mathcal{P}' \equiv \{\tau_{k_n} \,:\, n \in \mathbb{N}\} = \{\tau_j \,:\, j \in \mathbb{N},\, c_j = 1\}$$

is a Poisson process with rate $W$. Moreover, the sequence $\{e_{k_n}\}_{n\in\mathbb{N}} = \{e_j\}_{j\in\mathbb{N}, c_j=1}$ is i.i.d. with $\mathbb{P}(e_1' = e) = w_e/W$ and independent from $\mathcal{P}'$. One can check that the "new" maps $I_{(t,s]}$ defined here can be obtained by replacing the $\mathcal{P}$ and $\{e_n\}_n$ with $\mathcal{P}'$ and $\{e_{k_n}\}_{n\in\mathbb{N}}$ (respectively) in the construction of Section 3.2, and this cannot change the distribution of $I_{(t,s]}$. $\square$

## 5.2 The chameleon process

The *chameleon process* is built on top of the modified graphical construction. The definition of the process will depend on a parameter $T > 0$ which we call the *phase length*, for reasons that will become clear later on.

Given $\mathbf{y} \in (V)_{k-1}$, let $\mathbf{O}(\mathbf{y}) \equiv \{\mathbf{y}(1), \ldots, \mathbf{y}(k-1)\}$ denote the set of vertices that "occupied" by the coordinates of $\mathbf{y}$. The chameleon process will be a continuous-time, time-inhomogeneous Markov chain with state space:

$$\mathcal{C}_k(V) \equiv \{(\mathbf{z}, R, P, W) \,:\, \mathbf{z} \in (V)_{k-1};\ \text{the sets } \mathbf{O}(\mathbf{z}), R, P, W \text{ partition } V\}. \tag{5.2.1}$$

Notice that we do allow any of the $R, P, W$ to be empty in the above definition. For a given $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$, it will be convenient to refer to the vertices in the sets $\mathbf{O}(\mathbf{z}), R, P, W$ as *black*, *red*, *pink* and *white* (respectively). Notice that any vertex $v \in V$ will belong to one of these color classes.

The evolution of the process from an initial state $(\mathbf{z}, R, P, W)$ will be denoted by $\{(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C)\}_{t\ge 0}$. We first present an informal description of the process. There will be two kinds of *phases* to the process.

- *Constant-color phases* correspond to time intervals of the form $((2k-2)T, (2k-1)T]$, for $k \in \mathbb{N}\backslash\{0\}$. At each time $t = \tau_n$ belonging to one of these intervals, the process will move from its state $(\mathbf{z}^C_{t_-}, R^C_{t_-}, P^C_{t_-}, W^C_{t_-})$ immediately prior to time $t$ to the state:

$$(\mathbf{z}^C_t, R^C_t, P^C_t, W^C_t) = \begin{cases} (f_{e_n}(\mathbf{z}^C_{t_-}), f_{e_n}(R^C_{t_-}), f_{e_n}(P^C_{t_-}), f_{e_n}(W^C_{t_-})) & \text{if } c_n = 1; \\ (\mathbf{z}^C_{t_-}, R^C_{t_-}, P^C_{t_-}, W^C_{t_-}) & \text{if } c_n = 0. \end{cases} \tag{5.2.2}$$

  That is, the states of the endpoints of $e_n$ are flipped iff $c_n = 1$.

- *Color-changing phases* correspond to time intervals of the form $((2k-1)T, 2kT]$, for $k \in \mathbb{N}\backslash\{0\}$. At each time $t = \tau_n$ belonging to one of these intervals, the same update rule in the previous item is used, *unless* one of the endpoints of $e_n$ is white (ie. belongs to $W^C_{t_-}$), the other endpoint is red (ie. belongs to $R^C_{t_-}$) and the other conditions for a PINKENING TIME are satisfied (cf. Box 5.2). In that case, the red and white endpoints of $e_n$, $r \in R^C_{t_-}$ and $w \in W^C_{t_-}$, both become pink:

$$(\mathbf{z}^C_t, R^C_t, P^C_t, W^C_t) = (\mathbf{z}^C_{t_-}, R^C_{t_-}\backslash\{r\}, P^C_{t_-} \cup \{r, w\}, W^C_{t_-}\backslash\{w\}). \tag{5.2.3}$$

  In addition, at each time $2kT$ – i.e. at the end of each color-changing phase –, if the number of pink particles is greater than or equal to *either* the number white or the number of red particles, a fair coin $d_k$ is flipped; if it comes heads, all pink particles become red; otherwise, all pink become white. We call this a *depinking* step. For later convenience, we will also have a depinking step at time 0 if the conditions for depinking are met at that time.

**Remark 5.2.1** *Technically, this process is* not *càdlàg, as it changes at time* 0*. However, one can see that it* is *càdlàg outside of this point, so we will continue to use our notational convention*

$$(\mathbf{z}^C_{t_-}, R^C_{t_-}, P^C_{t_-}, W^C_{t_-})$$

*for all* $t > 0$ *(cf. Notational convention 2.3.1).*

**Remark 5.2.2** *We briefly note that our chamaleon process is much more complicated than that of Morris [20]. In brief: his process does not have constant-color phases; does not have additional conditions for pinkening other than* 2*. in Box 5.2; and will depink right when the number of pink particles exceeds the minimum of red and white. The differences we introduced will be fundamental at key steps of our argument.*

More formally, let:

$$\{\eta_0 = 0 < \eta_1 < \eta_2 < \dots\} = \{\tau_n\}_n \cup \{2kT\}_{k\in\mathbb{N}}$$

**Box 5.1** Conditions for $\tau_n \equiv \eta_i$ to be a PINKENING TIME.

Let $(\mathbf{z}^C_{\eta_{i-1}}, R^C_{\eta_{i-1}}, P^C_{\eta_{i-1}}, W^C_{\eta_{i-1}}) \equiv$ state of the process immediately before time $\tau_n$.
When $|R^C_{\eta_{i-1}}| \leq |W^C_{\eta_{i-1}}|$ (i.e. more white than red), the conditions are:

1. the process is in a color-changing phase, i.e. $\tau_n \in ((2k-1)T, 2kT]$ for some $k \in \mathbb{N}\backslash\{0\}$;
2. $e_n$ has one white endpoint and one red endpoint, that is $e_n = \{r_i, w_i\}$ with $r_i \in R^C_{\eta_{i-1}}$ red and $w_i \in W^C_{\eta_{i-1}}$ white;
3. there is no previous time in the same color-changing phase where an edge touched $r_i$; i.e. for all $m \leq n$ with $\tau_m \in ((\ell-1)T, \tau_n]$ and $r_i \in e_m$.
4. $|P^C_{\eta_{i-1}}| < |R^C_{\eta_{i-1}}|$

When $|R^C_{\eta_{i-1}}| > |W^C_{\eta_{i-1}}|$, substitute $w_i$ for $r_i$ in condition 3 and $|W^C_{\eta_{i-1}}|$ for $|R^C_{\eta_{i-1}}|$ in condition 4.

be the increasing sequence comprising all times $\tau_n \in \mathcal{P}$ and all times $2kT$ with $k \in \mathbb{N}$. We also define $\eta_{-1} = -1$ for convenience. By definition, the chameleon process will only change states at the times $\eta_i$ with $i \geq 0$, ie. it will be constant in each time interval $[\eta_{i-1}, \eta_i)$ with $i \in \mathbb{N}$. The updates at the times $\eta_i$ are described as an algorithm in Box 5.2 below. Notice that the sequence $\{d_k\}_{k\in\mathbb{N}}$ is an iid uniform $\{0,1\}$-sequence that is independent of $\mathcal{P}$, $\{e_n\}_{n\in\mathbb{N}}$ and $\{c_n\}_{n\in\mathbb{N}}$. We note that we do not distinguish explicitly between constant-color and color changing phases in our "pseudo-code", but it is easy to check that it does indeed implement the above update rules.

## 5.3 Some basic properties

Before we proceed, we note some properties of our chameleon process that will be .

**Lemma 5.3.1** *Let $i_1, i_2, i_3, \ldots$ be defined by the relation $\eta_{i_k} = 2kT$. Let*

$$(\hat{\mathbf{z}}_0, \hat{R}_0, \hat{P}_0, \hat{W}_0) = (\mathbf{z}, R, P, W)$$

*denote the initial state of the process and*

$$(\hat{\mathbf{z}}_k, \hat{R}_k, \hat{P}_k, \hat{W}_k) = \textit{the value of } (\mathbf{z}^C_{2kT_-}, R^C_{2kT_-}, P^C_{2kT_-}, W^C_{2kT_-}) \ \ (k \in \mathbb{N}\backslash\{0\}).$$

*Then $\{(\hat{\mathbf{z}}_k, \hat{R}_k, \hat{P}_k, \hat{W}_k)\}_{k\in\mathbb{N}}$ is a discrete-time, time-homogeneous Markov chain. Moreover, if $D_j$ is the $j$-th depinking time of the process, then $\hat{D}_j \equiv D_j/2T$ is a stopping-time for this discrete-time Markov chain.*

**Box 5.2** Evolution of the chameleon process given initial state $(\mathbf{z}, R, P, W)$.

1: **set** $(\mathbf{z}^C_{\eta-1}, R^C_{\eta-1}, P^C_{\eta-1}, W^C_{\eta-1}) \equiv (\mathbf{z}, R, P, W)$.
2: **for** $i \geq 0$ **do**
3:  **if** $\eta_i = \tau_n$ for some $n \in \mathbb{N}$ **then**
4:  **if** $\eta_i = \tau_n$ is a PINKENING TIME (cf. Box 5.2) **then**
5:  $e_n = \{r_i, w_i\}$ with $r_i \in R^C_{\eta_{i-1}}$ red and $w_i \in W^C_{\eta_{i-1}}$ white.
6:  **set** $R^C_{\eta_i} \leftarrow R^C_{\eta_{i-1}} \backslash \{r_i\}$ ▷ $r_i$ is not red anymore
7:  **set** $W^C_{\eta_i} \leftarrow W^C_{\eta_{i-1}} \backslash \{w_i\}$ ▷ $w_i$ is not white anymore
8:  **set** $P^C_{\eta_i} \leftarrow P^C_{\eta_{i-1}} \cup \{r_i, w_i\}$ ▷ both $r_i, w_i$ become pink.
9:  **set** $\mathbf{z}^C_{\eta_i} \leftarrow \mathbf{z}^C_{\eta_{i-1}}$ ▷ $\mathbf{z}$ part stays the same.
10:  **else** ▷ regular update: flip endpoints of $e_n$ iff $c_n = 1$
11:  **if** $c_n = 1$ **then**
12:  **set** $R^C_{\eta_i} \leftarrow f_{e_n}(R^C_{\eta_{i-1}})$
13:  **set** $W^C_{\eta_i} \leftarrow f_{e_n}(W^C_{\eta_{i-1}})$
14:  **set** $P^C_{\eta_i} \leftarrow f_{e_n}(P^C_{\eta_{i-1}})$
15:  **set** $\mathbf{z}^C_{\eta_i} \leftarrow f_{e_n}(\mathbf{z}^C_{\eta_{i-1}})$
16:  **else** ▷ $c_n = 0$
17:  **set** $(\mathbf{z}^C_{\eta_i}, R^C_{\eta_i}, P^C_{\eta_i}, W^C_{\eta_i}) \leftarrow$(state at time $\eta_{i-1}$).
18:  **end if**
19:  **end if**
20:  **else** ▷ $\eta_i = 2kT$ with $k \in \mathbb{N}$, cf. line 4.
21:  **if** $|P^C_{\eta_{i-1}}| \geq \min\{|W^C_{\eta_{i-1}}|, |R^C_{\eta_{i-1}}|\}$ **then** ▷ $\ell T$ is a depinking time
22:  **set** $P^C_{\eta_i} \leftarrow \emptyset$ ▷ all pink become red or white (see below)
23:  **set** $\mathbf{z}^C_{\eta_i} \leftarrow \mathbf{z}^C_{\eta_{i-1}}$ ▷ $\mathbf{z}$ part stays the same.
24:  **if** $d_k = 1$ **then**
25:  **set** $R^C_{\eta_i} \leftarrow R^C_{\eta_{i-1}} \cup P^C_{\eta_{i-1}}$ ▷ all pink become red
26:  **set** $W^C_{\eta_i} \leftarrow W^C_{\eta_{i-1}}$
27:  **else** ▷ $d_k = 0$
28:  **set** $R^C_{\eta_i} \leftarrow \emptyset$
29:  **set** $W^C_{\eta_i} \leftarrow W^C_{\eta_{i-1}} \cup P^C_{\eta_{i-1}}$ ▷ all pink become white
30:  **end if**
31:  **else** ▷ $\ell T$ not a depinking time; no changes made.
32:  **set** $(\mathbf{z}^C_{\eta_i}, R^C_{\eta_i}, P^C_{\eta_i}, W^C_{\eta_i}) \leftarrow$(state at time $\eta_{i-1}$)
33:  **end if**
34:  **end if**
35: **end for**

*Proof:* Given $(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i)$ for $i \leq k-1$, the state $(\hat{\mathbf{z}}_k, \hat{R}_k, \hat{P}_k, \hat{W}_k)$ consists of:

1. a depinking step on $(\hat{\mathbf{z}}_{k-1}, \hat{R}_{k-1}, \hat{P}_{k-1}, \hat{W}_{k-1})$ if $|\hat{P}_{k-1}| \geq \max\{|\hat{R}_{k-1}|, |\hat{W}_{k-1}|\}$;
2. having the resulting state go through a constant-color phase and a color-changing phase, without the final depinking step.

It is easy to see that the randomness used in 1. and 2. is independent of the entire past of the process, as it corresponds to the points and markings of the Poisson process in $[2(k-1)T, 2kT)$, as well as the extra coins used for depinking. Moreover, the random operations performed do not depend on time. The Markovianity and time-homogeneity of the process under consideration easily follow. To prove the last assertion, it suffices to check that (setting $D_0 = 0$),

$$\forall j > 0,\ \frac{D_j}{2T} = \inf\left\{k > \frac{D_{j-1}}{2T}\ :\ |\hat{P}_k| \geq \max\{|\hat{R}_k|, |\hat{W}_k|\}\right\}$$

where we allow the inf to be $+\infty$ if the set is empty or $D_{j-1} = +\infty$. □

**Lemma 5.3.2** *Suppose $(\mathbf{z}^C_{2iT}, R^C_{2iT}, P^C_{2iT}, W^C_{2iT})$ is the state of the chameleon process at time $2iT$ (ie. at the beginning of a constant-color phase). Then:*

$$(\mathbf{z}^C_{(2i+1)T}, R^C_{(2i+1)T}, P^C_{(2i+1)T}, W^C_{(2i+1)T}) = (I(\mathbf{z}^C_{2iT}), I(R^C_{2iT}), I(P^C_{2iT}), I(W^C_{2iT}))$$

*where $I = I_{(2iT,(2i+1)T]}$ is the map defined in the modified graphical construction.*

*Proof:* Follows from inspection. □

## 5.4 The chameleon process and conditional distributions

We now explain the relationship between the chameleon process and conditional distributions.

**Notational convention 5.4.1** *$\mathbf{x} = (\mathbf{x}(1), \ldots, \mathbf{x}(k)) \in (V)_k$ is represented as a pair $(\mathbf{z}, x)$, where $\mathbf{z} = (\mathbf{x}(1), \ldots, \mathbf{x}(k-1)) \in (V)_{k-1}$ and $x = \mathbf{x}(k) \in V \backslash \mathbf{O}(\mathbf{z})$. (Notice that $\mathbf{x}^I_t = (\mathbf{z}^I_t, x^I_t)$ for all $t \geq 0$.)*

**Proposition 5.4.1 (Proven in Section A.1)** *Given an initial state $\mathbf{x} = (\mathbf{z}, x) \in (V)_k$ for $\mathsf{IP}(k, G)$, set $R = \{x\}$, $P = \emptyset$ and $W = V \backslash (\mathbf{O}(\mathbf{z}) \cup \{x\})$. Consider the interchange process $\{\mathbf{x}^I_t = (\mathbf{z}^I_t, x^I_t)\}_{t \geq 0}$ started from state $\mathbf{x}$ and the chameleon process $\{(\mathbf{z}^C_t, R^C_t, P^C_t, W^I_t)\}_{t \geq 0}$ started from configuration $(\mathbf{z}, R, P, W) \in C_k(V)$. Then:*

$$\forall t \geq 0,\ \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\ \mathbb{P}\left(\mathbf{x}^I_t = \mathbf{b}\right) = \mathbb{E}\left[\mathsf{ink}_t(b)\mathbb{I}_{\{\mathbf{z}^C_t = \mathbf{c}\}}\right] \tag{5.4.1}$$

*where*

$$\mathsf{ink}_t(v) \equiv \mathbb{I}_{\{v \in R_t^C\}} + \frac{\mathbb{I}_{\{v \in P_t^C\}}}{2} \ (v \in V). \tag{5.4.2}$$

This result is our version for [20, Lemma 1], and the proof is similar; therefore, we have placed its proof in Section A.1 of the Appendix. Here we note that it is useful to think of $\mathsf{ink}_t(v)$ as the amount of "red ink" at vertex $v \in V$: a red vertex has one unit of red ink, a pink vertex has half a unit, and black or white vertices have no ink. We will see below that the *total* amount of red ink in the system determines the rate of convergence to equilibrium of $\mathsf{IP}(k, G)$.

# 6 Main proofs, via the chameleon process

We now present the proofs of Lemma 1.4.2 and Theorem 1.1.1, modulo several Lemmas about the chameleon process that we will prove later. We then outline the remainder of the paper.

## 6.1 Proof of Theorem 1.1.1 assuming Lemma 1.4.2

*Proof:* We will let $C = 10000K$ where $K$ is the constant appearing in Lemma 1.4.2. The combination of that with Lemma 1.4.1 gives:

$$\mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon) \leq C\, \mathsf{T}_{\mathsf{RW}(G)}(1/4)\ \ln(|V|/\epsilon) \text{ if } \epsilon \in (0, 1/2) \text{ and } k \leq |V|/2.$$

The contraction principle [2] – or direct inspection of the graphical construction – gives:

$$\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \leq \mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon) \leq C\, \mathsf{T}_{\mathsf{RW}(G)}(1/4)\ \ln(|V|/\epsilon) \text{ if } \epsilon \in (0, 1/2) \text{ and } k \leq |V|/2.$$

However, $\mathsf{EX}(k, G)$ and $\mathsf{EX}(|V| - k, G)$ are essentially the same process with the roles of empty and occupied sites reversed. In particular, $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) = \mathsf{T}_{\mathsf{EX}(|V|-k,G)}(\epsilon)$ for all $k, \epsilon$, and this finishes the proof. □

## 6.2 Proof of Lemma 1.4.2

*Proof:* [of Lemma 1.4.2] We assume we have defined a chameleon process over $\mathcal{C}_k(V)$ as in Section 5.2. We will take the notation and definitions from that section for granted. We also define:

$$\mathsf{ink}_t \equiv \sum_{v \in V} \mathsf{ink}_t(v) = |R_t^C| + \frac{|P_t^C|}{2}. \ \ (t \geq 0) \tag{6.2.1}$$

We note for later reference that

$$\mathsf{ink}_t \equiv \sum_{v \in V \setminus \mathbf{O}(\mathbf{z}_t^I)} \mathsf{ink}_t(v) \tag{6.2.2}$$

since the vertices in $\mathbf{O}(\mathbf{z}_t^I)$ have zero red ink.

We have argued in Proposition 5.4.1 that the distribution of $\mathsf{IP}(k, G)$ started from $\mathbf{x} = (\mathbf{z}, x) \in (V)_k$ corresponds to a chameleon process started from $(\mathbf{z}, \{x\}, \emptyset, V \setminus (\mathbf{O}(\mathbf{z}) \cup \{x\}))$. Letting $\mathsf{ink}_t^{\mathbf{x}}$ denote the value of $\mathsf{ink}_t$ in that chameleon process, we will show that:

**Lemma 6.2.1 (Proven in Section 8.2)** *The following inequality holds for all $1 \leq k \leq |V| - 1$:*

$$\sup_{\mathbf{x} \in (V)_k} d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathrm{Unif}((V)_k)) \leq 10k \mid \sup_{\mathbf{x} \in (V)_k} \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{x}}}{|V| - k + 1} \,\middle|\, \mathsf{Fill}\right]$$

*where*

$$\mathsf{Fill} \equiv \left\{ \lim_{t \to +\infty} \mathsf{ink}_t^{\mathbf{x}} = |V| - k + 1 \right\}.$$

The main goal is to bound the expected value in the RHS of the inequality in Lemma 6.2.1. Fix some $\mathbf{x} \in (V)_k$ and let $(\mathbf{z}, R, P, W) = (\mathbf{z}, \{x\}, \emptyset, V \setminus (\mathbf{O}(\mathbf{z}) \cup \{x\}))$ be the initial state corresponding to $\mathbf{x}$ in the sense of Proposition 5.4.1. Let $D_j(\mathbf{x})$ denote the $j$-th depinking time for this process, i.e. the $j$-th time $\eta_i = 2kT$, $k \in \mathbb{N}$, such that the **if** clause in Box 5.2 is satisfied. Also set $\widehat{\mathsf{ink}}_j^{\mathbf{x}} \equiv \mathsf{ink}_{D_j(\mathbf{x})}^{\mathbf{x}}$ for this process. We will show in Proposition 7.1.4 that there are infinitely many depinking times. The definition of the chameleon process implies that $\mathsf{ink}_t^{\mathbf{x}}$ can only change at depinking times, hence for any $t \geq 0$ $\mathsf{ink}_t^{\mathbf{x}} = 1$ if $t < D_1(\mathbf{x})$ and $\mathsf{ink}_t^{\mathbf{x}} = \widehat{\mathsf{ink}}_j^{\mathbf{x}}$ if $D_j(\mathbf{x}) \leq t < D_{j+1}(\mathbf{x})$ for some $j$. We deduce that:

$$1 - \frac{\mathsf{ink}_t^{\mathbf{x}}}{|V| - k + 1} \leq \sup_{m \geq j} \left(1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V| - k + 1}\right) + \mathbb{I}_{\{D_j(\mathbf{x}) > t\}} \leq \sum_{m \geq j} \left(1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V| - k + 1}\right) + \mathbb{I}_{\{D_j(\mathbf{x}) > t\}}.$$

Taking expectations, we see that the RHS of the inequality in Lemma 6.2.1 is at most:

$$10k \sup_{\mathbf{x} \in (V)_k} \left\{ \sum_{m \geq j} \mathbb{E}\left[1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V| - k + 1} \,\middle|\, \mathsf{Fill}\right] + \mathbb{P}\left(D_j(\mathbf{x}) \geq t \mid \mathsf{Fill}\right) \right\} \tag{6.2.3}$$

A simple (but technical) proposition will take care of the first term.

**Proposition 6.2.1 (Proven in Section 7.2)** *For all $\ell \geq 1$ and $\mathbf{x} \in (V)_k$,*

$$\mathbb{E}\left[1 - \frac{\widehat{\mathsf{ink}}_\ell^{\mathbf{x}}}{|V| - k + 1} \mid \mathsf{Fill}\right] \leq \sqrt{|V| - k + 1} \left(\frac{71}{72}\right)^\ell.$$

We thus have:

$$10k \sup_{\mathbf{x}\in(V)_k} \left\{ \sum_{m\geq j} \mathbb{E}\left[ 1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V|-k+1} \middle| \mathsf{Fill} \right] + \mathbb{P}\left(D_j(\mathbf{x}) \geq t \mid \mathsf{Fill}\right) \right\}$$
$$\leq C_2 \left|V\right|^{3/2} e^{-c_1 j} + 10k \sup_{\mathbf{x}\in(V)_k} \mathbb{P}\left(D_j(\mathbf{x}) \geq t \mid \mathsf{Fill}\right), \quad (6.2.4)$$

where $c_1 = \ln(72/71) > 0$ and $C_2 = 720$ are universal constants.

Bounding $\mathbb{P}\left(D_j(\mathbf{x}) \geq t \mid \mathsf{Fill}\right)$ is the key step in the proof. Up to now all of our results have been valid for all values of $k, |V|$ and of the phase length parameter $T > 0$. The next Lemma will require restrictions on these values.

**Lemma 6.2.2 (Proven in Section 9.3)** *There exist universal constants $C_2, K_2 > 0$ such that, if $|V| \geq 300$, $T \geq C_2\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$ and $k/|V| \leq 1/2$, then:*

$$\forall \mathbf{x} \in (V)_k,\, \forall j \in \mathbb{N} : \mathbb{E}\left[ e^{\frac{D_j(\mathbf{x})}{K_2 T}} | \mathsf{Fill} \right] \leq e^j.$$

If $|V| \geq 300$ Markov's inequality allows one to deduce that, for yet another universal constant $L_2 \equiv C_2 K_2$

$$\mathbb{P}\left(D_j(\mathbf{x}) > t | \mathsf{Fill}\right) \leq e^{j - \frac{t}{L_2\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)}}.$$

Plugging this into (6.2.4) and Lemma 6.2.1, we obtain:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I]), \mathrm{Unif}((V)_k)) \leq C_1 \left|V\right|^{3/2} e^{-c_1 j} + 10|V|\, e^{j - \frac{t}{L_2\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)}}.$$

Since this inequality holds for all $j$, we can take

$$j = \left\lfloor \frac{t}{2L_2\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)} \right\rfloor$$

and obtain:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathrm{Unif}((V)_k)) \leq K_0 \left|V\right|^{3/2} e^{-\frac{t}{K_1\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)}}$$

with $K_0, K_1 > 0$ universal. Comparing with the definition of mixing time in (1.1.1) and noting that $\mathrm{Unif}((V)_k)$ is stationary for $\mathsf{IP}(k,G)$ finishes the proof in the case $|V| \geq 300$.

The case $|V| < 300$ – ie. $|V|$ bounded by a universal constant – can be dealt with in several ways. For example, one may use the result of Caputo et al. [7] for the spectral gap of $\mathsf{IP}(k,G)$ together with the standard *lower* bound for $\mathsf{T}_{\mathsf{RW}(G)}(1/4)$ in terms of the spectral gap and the usual *upper bound* for $\mathsf{T}_{\mathsf{IP}(k,G)}(\epsilon)$ in terms of its spectral gap (see eg.

[17] for these standard bounds). Alternatively, one may use the analysis of Aldous and Fill (see Remark 4.2.1) together with the inequality

$$\mathbb{P}\left(M(\mathbf{x}) > 2k\mathsf{T}_{\mathsf{RW}(G)}(1/4)\right) \leq \left(1 - \frac{1}{4|V|}\right)^k,$$

which one can prove via Proposition 2.3.2 and a few simple calculations. □

### 6.3 Outline of the missing steps

We now summarize the main steps left in the proof.

1. In Section 7 we collect several facts about the quantity $\mathsf{ink}$. The proof of Proposition 6.2.1 on the decay of $\mathbb{E}\left[1 - \widehat{\mathsf{ink}}_\ell^{\mathbf{x}}/(|V| - k + 1) \mid \mathsf{Fill}\right]$ is presented in Section 7.2.

2. Section 8 contains the proof of Lemma 6.2.1, which is based on an auxiliary result on conditional distributions (Lemma 8.1.1).

3. Section 9 bounds the right tail of the first depinking time in a chameleon process, and then uses this to bound the exponential moment of the $j$-th depinking time. This leads to the key Lemma 6.2.2, proven in Section 9.3.

## 7 A miscellany of facts on ink

In this section we consider several of the facts that we will need in relation to the quantity $\mathsf{ink}_t$ introduced in (6.2.1). We will use the same notation introduced in the proof of Lemma 1.4.2 (cf. Section 6.2):

1. $\mathbf{x} \in (V)_k$ is some fixed state;

2. $(\mathbf{z}, R, P, W) = (\mathbf{z}, \{x\}, \emptyset, V \backslash (\mathbf{O}(\mathbf{z}) \cup \{x\})) \in \mathcal{C}_k(V)$ is the initial state corresponding to $\mathbf{x}$ in the sense of Proposition 5.4.1;

3. $\mathsf{ink}_t^{\mathbf{x}}$ is the total amount of ink in $(z_t^C, R_t^C, P_t^C, W_t^C)$ (with the above initial state);

4. $D_j(\mathbf{x})$ is the $j$-th depinking time for this process, i.e. the $j$-th time $\eta_i = 2kT$, $k \in \mathbb{N}$, such that the **if** clause in Box 5.2 is satisfied (also set $D_0(\mathbf{x}) = 0$);

5. finally, $\widehat{\mathsf{ink}}_j^{\mathbf{x}} \equiv \mathsf{ink}_{D_j(\mathbf{x})}^{\mathbf{x}}$.

We will mostly omit $\mathbf{x}$ from the notation in what follows.

## 7.1 Depinking times and the amount of ink

We observe that, in principle, the total number $J \geq 0$ of number of depinking times could be finite. Notice also that the $D_j$ are stopping times for the chameleon process.

**Proposition 7.1.1** *Suppose $j+1 \leq J$ (i.e. there are at least $j+1$ depinkings). Then $\widehat{\mathsf{ink}}_{j+1} \in \{\widehat{\mathsf{ink}}_j + \Delta(\widehat{\mathsf{ink}}_j), \widehat{\mathsf{ink}}_j - \Delta(\widehat{\mathsf{ink}}_j)\}$, where:*

$$\Delta(r) \equiv \left\lceil \frac{\min\{r, |V| - k + 1 - r\}}{3} \right\rceil \quad (r \in \mathbb{N}). \tag{7.1.1}$$

*Proof:* Lines 21 - 30 in Box 5.2 show that there are no pink particles left in the system after depinking is performed. In particular, this implies that $\widehat{\mathsf{ink}}_j \equiv \mathsf{ink}_{D_j} = |R^C_{D_j}|$. Now recall that:

$$(\mathbf{O}(\mathbf{z}^C_{D_j}), R^C_{D_j}, P^C_{D_j}, W^C_{D_j}) \text{ partition V},$$

where $\mathbf{O}(\mathbf{z}^C_{D_j})$ has $k-1$ elements, $R^C_{D_j}$ has $\widehat{\mathsf{ink}}_j$ elements and $P^C_{D_j}$ is empty. It follows that $W^C_{D_j}$ has $|V| - k + 1 - \widehat{\mathsf{ink}}_j$ elements, i.e. there are $\widehat{\mathsf{ink}}_j$ red and $|V| - k + 1 - \widehat{\mathsf{ink}}_j$ white particles at time $D_j$.

At each subsequent pinkening time, the number of red and white particles decrease by 1 each and the number of pink particles increases by 2. Thus after $p$ pinkening steps, there are

$$\widehat{\mathsf{ink}}_j - p \text{ red}, 2p \text{ pink, and } n - k - 1 - \widehat{\mathsf{ink}}_j - p \text{ white particles.}$$

The fourth condition for a pinkening time (cf. Box 5.2) implies that, once $2p \geq \widehat{\mathsf{ink}}_j - p$ or $2p \geq n - k - 1 - \widehat{\mathsf{ink}}_j - p$, no pinkenings can take place until there is an extra depinking. On the other hand, a depinking step will only take place when the number of pink particles is at least as large as the number of white or red particles.

This shows that, if there is an extra depinking, the number of $p$ particles immediately prior to that time is the *smallest* integer $p$ satisfying $2p \geq \widehat{\mathsf{ink}}_j - p$ or $2p \geq n - k - 1 - \widehat{\mathsf{ink}}_j - p$, which is $p = \Delta(\widehat{\mathsf{ink}}_j)$ for $\Delta$ defined in (7.1.1). We deduce:

> Assume that there is a depinking step after time $D_j$. Then the number of pink particles immediately prior to time $D_{j+1}$ is precisely $\Delta(\widehat{\mathsf{ink}}_j)$.

We have *assumed* that $j+1 \leq J$, in which case there *must* be at least one more depinking. At that time, either all pink particles become white, or they all become red. These possibilities corresponds to $\widehat{\mathsf{ink}}_{j+1} = \widehat{\mathsf{ink}}_j - \Delta(\widehat{\mathsf{ink}}_j)$ or $\widehat{\mathsf{ink}}_{j+1} = \widehat{\mathsf{ink}}_j + \Delta(\widehat{\mathsf{ink}}_j)$, respectively. □

**Proposition 7.1.2** $0 \leq \widehat{\mathsf{ink}}_j \leq |V| - k + 1$ *for all* $j \leq J$.

*Proof:* Follows from inspection of the formula (7.1.1) for $\Delta$. □

**Proposition 7.1.3** *Suppose that at a given stopping time $\sigma \geq 0$ for the chameleon process, $|P^C_\sigma| < \min\{|R^C_\sigma|, |W^C_\sigma|\}$. Then there almost surely is a time $t > \sigma$ when a pinkening step will be performed.*

*Proof:* We will show that there exists a universal $\delta > 0$ (independent of $\sigma$) and another stopping time $\sigma' > \sigma$ such that, conditionally on the process up to time $\sigma$, there is a probability $> \delta$ that there is a pinkening step in the time interval $(\sigma, \sigma']$. Using induction on this reasoning will give the desired result.

Let $t_0$ be another stopping time, defined as the smallest time of the form $(2k-1)T$, $k \in \mathbb{N}$, that is greater than $\sigma$. If there is a pinkening in the interval $(\sigma, t_0]$, we take $\sigma' = t_0$. Otherwise, there were no changes to the number of red, white and pink particles; in particular:

$$\min\{|R^C_{t_0}|, |W^C_{t_0}|\} > |P^C_{t_0}| \geq 0$$

and we can choose $r \in R^C_{t_0}$ and $w \in W^C_{t_0}$ in some consistent manner. Recall that the time interval $((2k-1)T, 2kT]$ is a constant-color phase and the process changes by a sequence of $f_{e_n}$ operations, at the times $\tau_n \in ((k-1)T, kT]$ with $c_n = 1$. Let $\tilde{I}$ denote the composition of these maps; it is easy to see that, conditionally on the process up to time $(k-1)T$, $\tilde{I}$ has the same law as the $I_{(0,T]}$ map appearing in the modified graphical construction of $\mathsf{IP}(k, G)$. In particular, given $e = \{a, b\} \in E$, there is a probability $\delta_1 > 0$ that $\tilde{I}(r) = a$ and $\tilde{I}(w) = b$, i.e., vertices $a$ and $b$ are red and white (respectively) at time $2kT$. Now consider the *first* time $\tau_m > 2kT$. Conditioned on everything else, there is a positive probability $\delta_2 > 0$ that $\tau_m \leq (2k+1)T$ $e_m = e = \{a, b\}$. We claim that, in this case, pinkening *must occur* at time $\tau_m$. Let us check the four conditions for $\tau_m$ to be a PINKENING TIME (cf. Box 5.2).

1. $\tau_m \in ((2k+1)T, (2k+2)T]$ with $\ell = k+1$ even.

2. the endpoints of $e_m$ are red and white.

3. $|P^C_{\tau_m}| < \min\{|R^C_{\tau_m}|, |W^C_{\tau_m}|\}$ because there was no pinkening in the time interval between $\sigma$ and $(k-1)T$ (by assumption) and there can be none in the interval $(2kT, (2k+1)T]$; so the number of pink, red and white particles cannot change. [Notice that depinkings also cannot occur, as the condition in line 21 of Box 5.2 cannot be verified.]

4. there cannot be a previous time in $((2k+1)T, \tau_m]$ where either $a$ or $b$ were touched by an edge, since $\tau_m$ is the *first* point of the Poisson process after time $kT$.

We deduce in this case that a pinkening will have occurred by time $(k+1)T$. Thus we can take $\sigma' = (k+1)T$ in this case and obtain that the overall conditional probability of a pinkening in the time interval $(\sigma, \sigma']$ is $\geq \delta_1\delta_2$. $\square$

**Proposition 7.1.4** *The number $J$ of depinking times is almost surely infinite. Moreover, once $\widehat{\mathsf{ink}}_j \in \{0, |V| - k - 1\}$, then $\widehat{\mathsf{ink}}_{j+k} = \widehat{\mathsf{ink}}_j$ for all $k \in \mathbb{N}$.*

**Remark 7.1.1** *Notice that this only works because our* definition *of a depinking time allows for "trivial depinking times" where there are no white and pink particles, or no red and pink particles.*

*Proof:* We first show that for any $j \in \mathbb{N} \cup \{0\}$ there almost surely are at least $j$ times that the **if** clause in line 21 of Box 5.2 is satisfied. This is clearly true for $j = 0$. Suppose it holds for some $j = j_0$. We consider three possibilities.

1. $\widehat{\mathsf{ink}}_{j_0} = \mathsf{ink}_{D_{j_0}} = |V| - k - 1$, i.e. there are $|V| - k - 1$ red particles at time $D_{j_0}$. Since there are $k-1$ black particles, and pink, white, black and red partition the vertex set $V$, there cannot be any white or pink particles in the system at this time. One can check that pinkening will not be performed again in this case, at least until the next potential depinking time, i.e. the smallest time $2kT > D_{j_0}$ with $k \in \mathbb{N}$. In particular, at this potential depinking time the number of pink and red particles will be 0, so the condition for the **if** in line 21 of Box 5.2 will be trivially satisfied. Thus $2kT = D_{j_0+1}$: there is a depinking time after $D_{j_0}$, even though there are no pink particles left.

2. $\widehat{\mathsf{ink}}_{j_0} = \mathsf{ink}_{D_{j_0}} = 0$, i.e. there are no red particles at time $D_{j_0}$. In this case one can show that there are no red or pink particles at time $D_{j_0}$. The same reasoning applied above proves that $D_{j_0+1}$ exists and is finite, even though there are no pink particles left.

3. $\widehat{\mathsf{ink}}_{j_0} \notin \{0, |V| - k + 1\}$. Notice once again that there are no pink particles at time $D_{j_0}$; moreover, there is a positive number of both red and white particles. In this case, iterated use of Proposition 7.1.3 shows that there will be more and more pinkenings until $|P_t^C| \geq \min\{|R_t^C|, |W_t^C|\}$ for some $t$. The first time $2kT$ after this $t$ must be a depinking step; i.e. $D_{j_0+1} = 2kT$ for this $k$.

We have shown that the number of depinking steps is a.s. lower bounded by $j = j_0 + 1$, starting from the analogous assertion for $j = j_0$. Induction implies the first assertion of the Proposition. The second assertion follows from inspection of cases 1 and 2 above. □

**Proposition 7.1.5** *Condition on the values of $\widehat{\mathsf{ink}}_1 = i_1, \ldots, \widehat{\mathsf{ink}}_j = i_j$ and assume they are all between 1 and $|V| - k - 2$. Then:*

$$\mathbb{P}\left(\widehat{\mathsf{ink}}_{j+1} = i_j - \Delta(i_j) \mid (\widehat{\mathsf{ink}}_r)_{r \leq j} = (i_r)_{r \leq j}\right) = \\ = \mathbb{P}\left(\widehat{\mathsf{ink}}_{j+1} = i_j + \Delta(i_j) \mid (\widehat{\mathsf{ink}}_r)_{r \leq j} = (i_r)_{r \leq j}\right) = 1/2. \quad (7.1.2)$$

*Proof:* This follows from the following facts: there are infinitely many depinking times (cf. Proposition 7.1.4); the increments of the process are always $\pm\Delta(\widehat{\mathsf{ink}}_j)$ (cf. Proposition 7.1.1); and that these two possibilities are equally likely, given the previous history of the process (cf. lines 21-30 in Box 5.2 and the fact that $d_k$ is independent of the past, which can be easily checked). $\square$

The next lemma summarizes the above sequence or propositions and adds a useful remark.

**Lemma 7.1.1** *The sequence $\{\widehat{\mathsf{ink}}_j\}_{j\geq 0}$ is a Markov chain with initial state $\widehat{\mathsf{ink}}_0 = 1$, absorbing states at $0$ and $|V|-k+1$ and transition probabilities given by:*

$$p(a,b) \equiv \frac{1}{2}\left(\mathbb{I}_{\{b=a+\Delta(a)\}} + \mathbb{I}_{\{b=a-\Delta(a)\}}\right) \quad (a,b \in \{0,1,2,\ldots,|V|-k+1\}). \tag{7.1.3}$$

*Moreover, it is almost surely absorbed in finite time in either $0$ or $|V|-k+1$. Finally, the event:*

$$\mathsf{Fill} \equiv \left\{\lim_{j\to+\infty} \widehat{\mathsf{ink}}_j = |V|-k+1\right\} \tag{7.1.4}$$

*has probability $1/(|V|-k+1)$.*

**Remark 7.1.2** *The event* $\mathsf{Fill}$ *corresponds to the number of red particles converging to $|V|-k+1$, i.e. that there are only black and red particles at all large enough times, or, equivalently, to red ink filling up all available space. Notice that we can rewrite:*

$$\mathsf{Fill} \equiv \left\{\lim_{t\to+\infty} \mathsf{ink}_t = |V|-k+1\right\},$$

*which is the form that appears in the proof of Lemma 1.4.1.*

*Proof:* [of Lemma 7.1.1] The first sentence is obvious given the sequence of Propositions; only notice that $p(a,a)=1$ if $a \in \{0,|V|-k+1\}$. We omit the trivial proof of the next assertion, which implies $\widehat{\mathsf{ink}}_\infty \equiv \lim_{j\to+\infty}\widehat{\mathsf{ink}}_j \in \{0,|V|-k+1\}$.

Now notice that the increments of $\widehat{\mathsf{ink}}_j$ are unbiased; that implies that this process is also a martingale. We thus have:

$$\begin{aligned}
\mathbb{P}(\mathsf{Fill}) &= \mathbb{P}\left(\widehat{\mathsf{ink}}_\infty = |V|-k+1\right) \\
(\widehat{\mathsf{ink}}_\infty \in \{0,|V|-k+1\}) &= \frac{\mathbb{E}\left[\widehat{\mathsf{ink}}_\infty\right]}{|V|-k+1} \\
(\{\widehat{\mathsf{ink}}_j\}_{j\in\mathbb{N}} \text{ bounded, cf. Prop. 7.1.2}) &= \frac{\lim_{j\to+\infty}\mathbb{E}\left[\widehat{\mathsf{ink}}_j\right]}{|V|-k+1} \\
(\text{mart. property} + \widehat{\mathsf{ink}}_0 = 1) &= \frac{\mathbb{E}\left[\widehat{\mathsf{ink}}_0\right]}{|V|-k+1} = \frac{1}{|V|-k+1}.
\end{aligned}$$

□

We will need one final lemma before we proceed.

**Lemma 7.1.2** *For all* $\mathbf{b} \in (V)_{k-1}$ *and* $t \geq 0$,

$$\mathbb{P}\left(\{\mathbf{z}_t^C = \mathbf{b}\} \cap \mathsf{Fill}\right) = \frac{\mathbb{P}\left(\mathbf{z}_t^C = \mathbf{b}\right)}{|V| - k + 1}.$$

*Proof:* This follows from the previous Lemma if we can show that $\mathsf{Fill}$ and $\mathbf{z}_t^C$ are independent. To see this, simply notice that $\mathsf{Fill}$ is entirely determined by the "coin flips" $d_k$ performed at the depinking times (cf. lines 21-30 of Box 5.2), whereas the value of $\mathbf{z}_t^C$ does not at all depend on these coin flips. [The last assertion is implicit in the proof of Proposition 5.4.1 and can also be checked directly.] □

## 7.2 The trajectory of $\widehat{\mathsf{ink}}_j$ given $\mathsf{Fill}$

We use the facts just prove to derive the technical estimate in Proposition 6.2.1 in the proof of Lemma 1.4.2 (cf. Section 6.2).

*Proof:* [of Proposition 6.2.1] We omit the superscript $\mathbf{w}$ in this proof. We note that

$$\widehat{\mathsf{ink}}_\infty = \lim_{j \to +\infty} \widehat{\mathsf{ink}}_j \in \{0, |V| - k + 1\}$$

by the results proven in Section 7.1.

Our first goal will be to show that, conditionally on $\mathsf{Fill}$, $\{\widehat{\mathsf{ink}}_j\}_{j \geq 0}$ is still a Markov chain. Repeating the steps of the proof of Lemma 7.1.1, we note that:

$$\begin{aligned}
\mathbb{P}\left(\mathsf{Fill} \mid (\widehat{\mathsf{ink}}_i)_{i \leq j}\right) &= \mathbb{P}\left(\widehat{\mathsf{ink}}_\infty = |V| - k + 1 \mid (\widehat{\mathsf{ink}}_i)_{i \leq j}\right) \\
(\mathsf{ink}_\infty \in \{0, |V| - k + 1\}) &= \frac{\mathbb{E}\left[\widehat{\mathsf{ink}}_\infty \mid (\widehat{\mathsf{ink}}_i)_{i \leq j}\right]}{|V| - k + 1} \\
(\{\widehat{\mathsf{ink}}_{j+k}\}_k \text{ uniformly bounded}) &= \lim_{k \to +\infty} \frac{\mathbb{E}\left[\widehat{\mathsf{ink}}_{j+k} \mid (\widehat{\mathsf{ink}}_i)_{i \leq j}\right]}{|V| - k + 1} \\
(\text{martingale property}) &= \frac{\widehat{\mathsf{ink}}_j}{|V| - k + 1}.
\end{aligned}$$

We deduce from Bayes' rule and the Markovian property for the (unconditional) $\widehat{\mathsf{ink}}_j$ that:

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^{j}\{\widehat{\mathsf{ink}}_i = a_i\} \mid \mathsf{Fill}\right) &= \left(\frac{\mathbb{P}\left(\bigcap_{i=1}^{j}\{\widehat{\mathsf{ink}}_i = a_i\}\right)}{\mathbb{P}\left(\mathsf{Fill}\right)}\right)\mathbb{P}\left(\mathsf{Fill} \mid \bigcap_{i=1}^{j+1}\{\widehat{\mathsf{ink}}_i = a_i\}\right) \\
\text{(previous formula + Lemma 7.1.1)} &= \mathbb{P}\left(\bigcap_{i=1}^{j}\{\widehat{\mathsf{ink}}_i = a_i\}\right) a_j \\
\text{(Markov property for } \widehat{\mathsf{ink}}_j) &= p(1,a_1)p(a_1,a_2)\dots p(a_{j-1},a_j)\, a_j \\
&= q(1,a_1)\dots q(a_{j-1},a_j)
\end{aligned}$$

where

$$q(a,b) = \frac{bp(a,b)}{a} \text{ if } a \neq 0.$$

Notice that, since $\widehat{\mathsf{ink}}_j$ does not visit 0 in the event $\mathsf{Fill}$, we do not need to define $q(a,b)$ for $a = 0$. We have shown:

**Proposition 7.2.1** *Conditionally on* $\mathsf{Fill}$*, the trajectory of* $\{\widehat{\mathsf{ink}}_j\}_{j\geq 0}$ *is that of a Markov chain in* $\{1,\dots,|V|-k+1\}$*, with transition rates* $q(a,b)$ *and started from* $\widehat{\mathsf{ink}}_0 = 1$*.*

**Remark 7.2.1** *Notice that this implies that the chameleon process conditioned on* $\mathsf{Fill}$ *is the same as the unconditional process, with the only difference that the coin flips* $d_k$ *performed at depinking times are biased towards* 1*. This remark will be useful in the proof of Lemma 6.2.2 in Section 9.3.*

Fr the remainder of the proof, we will use this Proposition to bound $1-\widehat{\mathsf{ink}}_\ell/(|V|-k+1)$. It turns out that another quantity is easier to bound. Set $I_\ell = \widehat{\mathsf{ink}}_\ell/(|V|-k+1)$ and:

$$Z_\ell \equiv \frac{\sqrt{\min\{1-I_\ell, I_\ell\}}}{I_\ell}.$$

Notice that conditionally on $\mathsf{Fill}$, $I_\ell > 0$ always, hence $Z_\ell$ is a.s. well defined for all $\ell$. Moreover, one can check that $1 - I_\ell \leq Z_\ell$ always. Therefore the Lemma will follow from the estimate:

$$\mathbb{E}^{\mathsf{Fill}}\left[Z_\ell\right] \leq (71/72)^\ell \sqrt{|V|-k+1},$$

where $\mathbb{E}^{\mathsf{Fill}}\left[\cdot\right]$ corresponds to an expectation with respect to the conditional distribution given $\mathsf{Fill}$. Since $Z_0 = \sqrt{|V|-k+1}$, the above estimate follows directly from the following *claim.*

**Claim 1**

$$\forall \ell \in \mathbb{N},\ \mathbb{E}^{\mathsf{Fill}}\left[Z_\ell\right] \leq (71/72)\,\mathbb{E}\left[Z_{\ell-1}\right].$$

Therefore, proving this Claim will finish the proof.

To prove the claim, we first note that for all $i$ $Z_i$ is a function of $i$, hence we can write:

$$\mathbb{E}^{\mathsf{Fill}}[Z_\ell] = \mathbb{E}^{\mathsf{Fill}}\left[\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1}\right] Z_{\ell-1}\mathbb{I}_{\{Z_{\ell-1}\neq 0\}} + Z_\ell \mathbb{I}_{\{Z_{\ell-1}=0\}}\right].$$

The second term inside the expectation in the RHS is 0. Indeed, if $Z_{\ell-1} = 0$, this means $\min\{1 - I_{\ell-1}, I_{\ell-1}\} = 0$, and since $I_{\ell-1} > 0$ a.s. (conditionally on $\mathsf{Fill}$), we have $1 - I_{\ell-1} = 0$ in that case, which is the same as $\widehat{\mathsf{ink}}_{\ell-1} = |V| - k + 1$. This implies $\widehat{\mathsf{ink}}_\ell = |V| - k + 1$ and $Z_\ell = 0$ as well.

We can thus rewrite:

$$\mathbb{E}^{\mathsf{Fill}}[Z_\ell] = \mathbb{E}^{\mathsf{Fill}}\left[\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1}\right] Z_{\ell-1}\mathbb{I}_{\{Z_{\ell-1}\neq 0\}}\right]. \tag{7.2.1}$$

We now bound the conditional expectation in the RHS. Notice that we may assume that $\widehat{\mathsf{ink}}_{\ell-1} = a$ with $r \neq 0$ – $\widehat{\mathsf{ink}}_{\ell-1} \neq 0$ a.s. under the conditioning – and $r < |V| - k + 1$ – as otherwise $\widehat{\mathsf{ink}}_{\ell-1} = |V| - k + 1$ and $Z_{\ell-1} = 0$. Thus we wish to bound:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right], 1 \leq r \leq n - k.$$

There is a better way to write this expression. If we note that:

$$\frac{Z_\ell}{Z_{\ell-1}} = \frac{\sqrt{\min\{1 - I_\ell, I_\ell\}}}{\sqrt{\min\{1 - I_{\ell-1}, I_{\ell-1}\}}} \times \frac{I_{\ell-1}}{I_\ell} = \frac{\sqrt{\min\{1 - I_\ell, I_\ell\}}}{\sqrt{\min\{1 - I_{\ell-1}, I_{\ell-1}\}}} \times \frac{\widehat{\mathsf{ink}}_{\ell-1}}{\widehat{\mathsf{ink}}_\ell}$$

and define $f(a) = \sqrt{\min\{a, |V| - k + 1 - a\}}$, we see that:

$$\begin{aligned}
\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] &= \mathbb{E}^{\mathsf{Fill}}\left[\frac{f(\widehat{\mathsf{ink}}_\ell)}{f(\widehat{\mathsf{ink}}_{\ell-1})} \times \frac{\widehat{\mathsf{ink}}_{\ell-1}}{\widehat{\mathsf{ink}}_\ell} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] \\
\text{(use Proposition 7.2.1)} &= \sum_s q(r,s)\,\frac{f(s)}{f(r)} \times \frac{r}{s} \\
\text{(use formula for } q(\cdot,\cdot\cdot)) &= \sum_s p(r,s)\frac{f(s)}{f(r)}
\end{aligned}$$

where $p(\cdot,\cdot\cdot)$ are the transition rates of the unconditional $\{\widehat{\mathsf{ink}}_j\}_{j\geq 0}$ process. Using the formula for these we obtain:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] = \frac{1}{2}\left(\frac{f(r + \Delta(r)) + f(r - \Delta(r))}{f(r)}\right). \tag{7.2.2}$$

Recall the formula for $\Delta(r)$, cf. Proposition 7.1.1:

$$\Delta(r) \equiv \left\lceil \frac{\min\{r, |V| - k + 1 - r\}}{3} \right\rceil$$

We now split the analysis of the RHS of this in two cases.

**Case 1:** $1 \leq r \leq (|V| - k + 1)/2$. In this case $f(r) = \sqrt{r}$ and $\Delta(r) = \lceil r/3 \rceil \geq r/3$. We use the upper bound $f(r \pm \Delta(r)) \leq \sqrt{r \pm \Delta(r)}$ to obtain:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] = \frac{1}{2}\left(\sqrt{1 - \frac{\Delta(r)}{r}} + \sqrt{1 + \frac{\Delta(r)}{r}}\right). \tag{7.2.3}$$

Recall the bound "$\sqrt{1-x} + \sqrt{1+x} \leq 2(1 - x^2/8)$", valid for all $0 \leq x \leq 1$; this can be checked by squaring both sides of the inequality. In our case, we apply this with $x = \Delta(r)/r \geq 1/3$ and deduce:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] = 1 - \frac{1}{8}\left(\frac{\Delta(r)}{r}\right)^2 \leq \frac{71}{72}. \tag{7.2.4}$$

**Case 2:** $(|V| - k + 1)/2 < r \leq n - k$. Set $r' = |V| - k + 1 - r$ and notice that $1 \leq r' \leq (|V| - k + 1)/2$, $f(r) = \sqrt{r'}$, $\Delta(r') = \Delta(r)$ and

$$f(r \pm \Delta(r)) \leq \sqrt{|V| - k + 1 - r \mp \Delta(r)} = \sqrt{r' \mp \Delta(r)} = \sqrt{r' \mp \Delta(r')}.$$

The same calculations performed in the first case lead to:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] = \frac{1}{2}\left(\sqrt{1 - \frac{\Delta(r')}{r'}} + \sqrt{1 + \frac{\Delta(r')}{r'}}\right) \leq \frac{71}{72}. \tag{7.2.5}$$

Thus we see that in both cases:

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \mid \widehat{\mathsf{ink}}_{\ell-1} = r\right] \leq \frac{71}{72}.$$

Plugging this into (7.2.1) gives:

$$\mathbb{E}^{\mathsf{Fill}}\left[Z_\ell\right] \leq \frac{71}{72}\,\mathbb{E}^{\mathsf{Fill}}\left[Z_{\ell-1}\mathbb{I}_{\{Z_{\ell-1} \neq 0\}}\right] = \frac{71}{72}\,\mathbb{E}^{\mathsf{Fill}}\left[Z_{\ell-1}\right],$$

which finishes the proof. $\square$

# 8 Convergence to stationarity in terms of ink

In this section we will prove Lemma 6.2.1, used in the proof of Lemma 1.4.2 (cf. Section 6.2), in which we show that the amount of ink in the system can be used to bound the distance to the stationary distribution. We start with a preliminary result on marginals.

## 8.1 The convergence to equilibrium of conditional distributions

We will again use Notational convention 5.4.1, whereby any $\mathbf{x} = (\mathbf{x}(1), \ldots, \mathbf{x}(k)) \in (V)_k$ is written as a pair $\mathbf{x} = (\mathbf{z}, x)$ with $\mathbf{z} = (\mathbf{x}(1), \ldots, \mathbf{x}(k-1))$ and $x = \mathbf{x}(k)$.

Let $\mathbf{x} = (\mathbf{z}, x) \in (V)_k$ and consider the $\mathsf{IP}(k, G)$ process $\{\mathbf{x}_t^I\}_{t\geq 0}$. Set $R = \{x\}$, $P = \emptyset$ and $W = V \backslash (\mathbf{O}(\mathbf{z}) \cup \{x\})$ and recall from Proposition 5.4.1 that the chameleon process $\{(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C)\}_{t\geq 0}$ satisfies:

$$\forall t \geq 0,\ \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\ \mathbb{P}\left(\mathbf{x}_t^I = \mathbf{b}\right) = \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\}} \mathsf{ink}_t^{\mathbf{x}}(b)\right] \tag{8.1.1}$$

where (as before) we use $\mathsf{ink}_t^{\mathbf{x}}(\cdot)$ to denote the amount of ink in this chameleon process corresponding to $\mathbf{x}$. The following lemma relates the *total* amount of ink in this process to the near-uniformity of the last coordinate of $bx_t^I$ given $\mathbf{z}_t^I$.

**Lemma 8.1.1** *For all* $\mathbf{x} \in (V)_k$,

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I]), \mathcal{L}[\tilde{\mathbf{x}}_t^I]) \leq \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{x}}}{|V| - k + 1} \,\middle|\, \mathsf{Fill}\right]$$

*where* $\mathsf{Fill}$ *is the event defined in Lemma 7.1.1 (see also Remark 7.1.2).*

*Proof:* One can deduce from the (modified) graphical construction in Section 5.1 that the conditional distribution of $\tilde{x}_t^I$ given $\mathbf{z}_t^I$ is uniform over $V \backslash \mathbf{O}(\mathbf{z}_t^I)$. Moreover, we have seen that $\mathbf{z}_t^I$ and $\mathbf{z}_t^C$ have the same distribution (cf. the proof of Proposition 5.4.1). We deduce that:

$$\forall t \geq 0,\ \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\ \mathbb{P}\left(\mathbf{z}_t^I = \mathbf{c}, \tilde{x}_t^I = b\right) = \frac{\mathbb{P}\left(\mathbf{z}_t^C = \mathbf{c}\right)}{|V| - k + 1} = \mathbb{P}\left(\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}\right)$$

where the last equality follows from Lemma 7.1.2. On the other hand, (8.1.1) implies:

$$\forall t \geq 0,\ \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\ \mathbb{P}\left(\mathbf{x}_t^I = \mathbf{b}\right) \geq \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}} \mathsf{ink}_t^{\mathbf{x}}(b)\right]. \tag{8.1.2}$$

We deduce that:

$$\begin{aligned}
\forall t \geq 0,\ \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\ &(\mathbb{P}\left(\mathbf{z}_t^I = \mathbf{c}, \tilde{x}_t^I = b\right) - \mathbb{P}\left(\mathbf{z}_t^I = \mathbf{c}, \tilde{x}_t^I = b\right))_+ \leq \\
&\leq \left(\mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}}(1 - \mathsf{ink}_t^{\mathbf{x}}(b))\right]\right)_+ = \\
&= \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}}(1 - \mathsf{ink}_t^{\mathbf{x}}(b))\right] \text{ since the integrand is } \geq 0.
\end{aligned}$$

Using formula (2.2.3) for $d_{\mathrm{TV}}(\cdot,\cdot\cdot)$ gives:

$$\begin{aligned}
d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\tilde{\mathbf{x}}_t^I]) &= \sum_{\mathbf{b}=(\mathbf{c},b)\in(V)_k} \left(\mathbb{P}\left(\mathbf{z}_t^I=\mathbf{c},\tilde{x}_t^I=b\right)-\mathbb{P}\left(\mathbf{z}_t^I=\mathbf{c},\tilde{x}_t^I=b\right)\right)_+ \\
\text{(use previous ineq.)} &\leq \sum_{\mathbf{b}=(\mathbf{c},b)\in(V)_k} \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C=\mathbf{c}\}\cap\mathsf{Fill}}(1-\mathsf{ink}_t^{\mathbf{x}}(b))\right] \\
&= \sum_{\mathbf{c}\in(V)_{k-1}} \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C=\mathbf{c}\}\cap\mathsf{Fill}} \sum_{b\in V\backslash\mathbf{O}(\mathbf{c})}(1-\mathsf{ink}_t^{\mathbf{x}}(b))\right] \\
\text{(sum over } b + (6.2.2)) &= \sum_{\mathbf{c}\in(V)_{k-1}} \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C=\mathbf{c}\}\cap\mathsf{Fill}}(|V|-k+1-\mathsf{ink}_t^{\mathbf{x}})\right] \\
\text{(sum over } \mathbf{c}) &= \mathbb{E}\left[\mathbb{I}_{\mathsf{Fill}}(|V|-k+1-\mathsf{ink}_t^{\mathbf{x}})\right] \\
\text{(apply Lemma 7.1.1)} &= 1-\frac{\mathbb{E}\left[\mathsf{ink}_t^{\mathbf{x}} \mid \mathsf{Fill}\right]}{|V|-k+1}.
\end{aligned}$$

□

## 8.2 Distance to the stationary distribution in terms of ink

*Proof:* [of Lemma 6.2.1] We will prove the following stronger inequality:

$$\sup_{\mathbf{x},\mathbf{y}\in(V)_k} d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\mathbf{y}_t^I]) \leq 10k \sup_{\mathbf{w}\in(V)_k} \mathbb{E}\left[1-\frac{\mathsf{ink}_t^{\mathbf{w}}}{|V|-k+1}\,\middle|\,\mathsf{Fill}\right], \tag{8.2.1}$$

which implies the Lemma by convexity.

Declare two states $\mathbf{u},\mathbf{v}\in(V)_k$ to be *adjacent* ($\mathbf{u}\sim\mathbf{v}$) if they differ at precisely one coordinate: i.e there exists an $i\in[k]$ with $\mathbf{u}(i)\neq\mathbf{v}(i)$ and $\mathbf{u}(r)=\mathbf{v}(r)$ for $r\in[k]\backslash\{i\}$. We first bound $d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\mathbf{y}_t^I])$ for *adjacent* $\mathbf{x}\sim\mathbf{y}$.

One can assume without loss of generality that $\mathbf{x}$ and $\mathbf{y}$ differ precisely at the $k$-th coordinate, as shuffling the coordinates before the process starts does not alter $d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\mathbf{y}_t^I])$. Using the notation from Section 5.4, this allows us to write $\mathbf{x}=(\mathbf{z},x)$ and $\mathbf{y}=(\mathbf{z},y)$ for $\mathbf{z}\in(V)_{k-1}$ and $x\in V\backslash\mathbf{O}(\mathbf{z})$. Defining $\tilde{\mathbf{x}}=(\mathbf{z},\tilde{x})$ as in Section 8.1 and $\tilde{\mathbf{y}}$ similarly, we see that $\mathcal{L}[\tilde{\mathbf{x}}]=\mathcal{L}[\tilde{\mathbf{y}}]$ and therefore $\mathcal{L}[\tilde{\mathbf{x}}_t^I]=\mathcal{L}[\tilde{\mathbf{y}}_t^I]$ for all $t\geq 0$. We deduce that:

$$\begin{aligned}
d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\mathbf{y}_t^I]) &\leq d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\tilde{\mathbf{x}}_t^I]) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_t^I],\mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
&\quad + d_{\mathrm{TV}}(\mathcal{L}[\tilde{\mathbf{x}}_t^I],\mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
\text{(3rd. term= 0)} &= d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I],\mathcal{L}[\tilde{\mathbf{x}}_t^I]) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_t^I],\mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
\text{(use Lemma 8.1.1)} &\leq \mathbb{E}\left[1-\frac{\mathsf{ink}_t^{\mathbf{x}}}{|V|-k+1}\,\middle|\,\mathsf{Fill}\right] + \mathbb{E}\left[1-\frac{\mathsf{ink}_t^{\mathbf{y}}}{|V|-k+1}\,\middle|\,\mathsf{Fill}\right] \\
&\leq 2\sup_{\mathbf{w}\in(V)_k} \mathbb{E}\left[1-\frac{\mathsf{ink}_t^{\mathbf{w}}}{|V|-k+1}\,\middle|\,\mathsf{Fill}\right] \qquad (8.2.2)
\end{aligned}$$

Now consider $\mathbf{x}, \mathbf{y} \in (V)_k$ arbitrary. One can find a sequence $\{\mathbf{x}[i]\}_{i=0}^r \subset (V)_k$ with $r \leq 10k$ and:

$$\mathbf{x}[0] = \mathbf{x} \sim \mathbf{x}[1] \sim \mathbf{x}[2] \sim \cdots \sim \mathbf{x}[r] = \mathbf{y}.$$

The triangle inequality gives:

$$d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\mathbf{y}_t^I]) = d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}[0]_t^I], \mathcal{L}[\mathbf{x}[r]_t^I]) \leq \sum_{i=1}^r d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}[i-1]_t^I], \mathcal{L}[\mathbf{x}[i]_t^I]).$$

Applying (8.2.2) to each adjacent pair $\mathbf{x}[i-1], \mathbf{x}[i]$ gives (8.2.1). $\square$

# 9 Depinkings are fast

The results in this section lead to the key Lemma 6.2.2. We will first need to consider what happens in the first two phases of the chameleon process (constant color/color-changing phases), started from an essentially arbitrary initial state. Our main goal will be to show that the number of red particles decreases in expectation by a constant factor. More precisely, we will show that:

**Lemma 9.0.1 (Proven in Section 9.1)** *Consider a suitably modified chameleon process (the modification will be described in the proof of the Lemma) on a graph $G = (V, E, \{w_e\}_{e \in E})$ with $|V| \geq 300$, started from an initial state $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$ with $|P| < |R| \leq |W|$ and $|W| \geq |V|/4$. If the phase length parameter $T$ satisfies:*

$$T \geq 20\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4),$$

*then:*

$$\mathbb{E}\left[|R_{2T_-}^C|\right] \leq (1-c)|R|$$

*where $c = 1/1000 > 0$ and $R_{2T_-}^C$ denotes the state of $R_t^C$ immediately prior to time $2T$.*

With this, it will not be hard to show that the first depinking time has an exponential moment.

**Lemma 9.0.2 (Proven in Section 9.2)** *Consider a chameleon process (without the modification in the previous Lemma) on a graph $G = (V, E, \{w_e\}_{e \in E})$ with $|V| \geq 300$, started from an initial state $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$ with $|P| = \emptyset$. There exists a universal constant $K > 0$ such that if the phase length parameter $T$ satisfies $T \geq \lceil 20\, \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$, the first depinking time $D_1$ of this process satisfies:*

$$\mathbb{E}\left[e^{\frac{D_1}{K\,T}}\right] \leq e.$$

**Box 9.1** New conditions for $\tau_n \equiv \eta_i$ to be a PINKENING TIME.

Let $(\mathbf{z}^C_{\eta_{i-1}}, R^C_{\eta_{i-1}}, P^C_{\eta_{i-1}}, W^C_{\eta_{i-1}}) \equiv$ state of the process immediately before time $\tau_n$.
When $|R^C_{\eta_{i-1}}| \leq |W^C_{\eta_{i-1}}|$ (i.e. more white than red), the conditions are:

1. the process is in a color-changing phase, i.e. $\tau_n \in ((2k-1)T, 2kT]$ for some $k \in \mathbb{N}\backslash\{0\}$;
2. $e_n$ has one white endpoint and one red endpoint, that is $e_n = \{r_i, w_i\}$ with $r_i \in R^C_{\eta_{i-1}}$ red and $w_i \in W^C_{\eta_{i-1}}$ white;
3. there is no previous time in the same color-changing phase where an edge touched $r_i$; i.e. for all $m \leq n$ with $\tau_m \in ((2k-1)T, \tau_n)$ and $r_i \in e_m$.

When $|R^C_{\eta_{i-1}}| > |W^C_{\eta_{i-1}}|$, substitute $w_i$ for $r_i$ in condition 3.

From this it will be a short step to prove that the $j$th depinking time has an expoential moment (cf. Section 9.3).

The proof of Lemma 9.0.1 will shed some light on some of the choices we have made in the construction of the chameleon process. We will emphasize this with several remarks in the proof.

## 9.1 Loss of red particles in the two first phases

*Proof:* [of Lemma 9.0.1] The "suitable modification" we make to the chameleon process is this: we drop condition 4. for a PINKENING TIME (cf. Box 5.2 in page 25). That is, we allow depinkings to happen even if $|P^C_{\eta_{i-1}}| \geq \min\{|W^C_{\eta_i}|, |R^C_{\eta_i}|\}$. The new conditions for a PINKENING TIME are given in Box 9.1 (cf. 43) below.

We also note that, given the assumptions of the Theorem, we will have $|R^C_t| \leq |W^C_t|$ for all $0 \leq t < 2T$, since the number of red and white particles both decrease by 1 at pinkening times and remain constant at all other times.

To continue, we recall the definition of the process (Box 5.2 in page 25) we know that there is no depinking right at the beginning of this process, since there are less pink particles than white or red ones. We also know that the time interval $(0, T]$ is a constant-color phase where black, red and white particles are simply moved around, whereas $(T, 2T]$ is a color-changing phase. More specifically, we recall from Lemma 5.3.2 that the state of the process at time $T$ is given by:

$$(\mathbf{z}^C_T, R^C_T, P^C_T, W^C_T) = (I(\mathbf{z}), I(R), I(P), I(W))$$

where $I = I_{(0,T]} = I_T$ is the map obtained from the modified chameleon construction in Section 5. We will need the following properties later on:

**Proposition 9.1.1 (Proven in Section 9.1.1)** *For all $(a,b) \in (V)_2$ and $S, L \subset V$ with $|L| \geq |V|/12$,*

$$\mathbb{P}\left((a,b) \in I(S) \times I(S)\right) \leq \mathbb{P}\left(a \in I(S)\right)\left(\frac{|S|}{|V|} + 2^{-10}\right).$$

$$\mathbb{P}\left((a,b) \in I(S) \times I(L)\right) \geq \frac{|S||L|}{|V|^2}\left(1 - 2^{-9}\right) \geq \frac{|S|}{13|V|}.$$

**Remark 9.1.1** *The intuitive meaning of this is that $(R_T^C, W_T^C)$ are close to uniform in terms of correlations of "pairs of particles" at the end of the constant-color phase, and this will only hold because $T = \Omega\left(\mathsf{T}_{\mathsf{IP}(2,G)}\right)$. Morris' original argument for $(\mathbb{Z}/L\mathbb{Z})^d$ could instead rely on good estimates for transition probabilities for single-particle random walks. We note that we need the negative correlation property in the proof of this proposition.*

In the time interval $(T, 2T)$, each time $T < \tau_m < 2T$ may be a pinkening time, depending on whether the pinkening conditions (cf. Box 5.2) are satisfied. We will nevertheless find it necessary to consider the maps:

$$\tilde{I}_t \equiv I_{(T,t]},\ T \leq t < 2T \text{ (cf. defn. in Section 5.1).)} \tag{9.1.1}$$

We emphasize that $\tilde{I}_t$ does *not* correspond directly to the evolution of the chameleon process in the time interval $(T, 2T)$, although we will use $\tilde{I}_t$ to understand the latter. Notice that Proposition 3.2.3 and Proposition 5.1.1 imply:

**Proposition 9.1.2 (Proof omitted)** *$\{\tilde{I}_t\}_{T<t<2T}$ is independent from $I$, and so are all the points of the Poisson process $\{\tau_n\}_n$ in the interval $(T, 2T)$ and all markings $e_n, c_n$ corresponding to these points.*

We need a new definition before we proceed. Let $a \in V$ be given. Let $\phi_a$ be the *first time* of the form $\tau_m$ with $T < \tau_m \leq 2T$ for which $a \in e_m$; if no such time exists, let $\phi_a = +\infty$. If $\phi_a < +\infty$, there exists a vertex $b \in V$ such that the edge $e_m$ just mentioned has $a = \tilde{I}_{\phi_a -}(a)$ and $\tilde{I}_{\phi_a -}(b)$ as endpoints immediately prior to time $\phi_a$. We set $F_a \equiv b$ in that case, or $F_a \equiv *$ if $\phi_a = +\infty$. The following simple claim is essential to what follows.

**Claim 2 (proven in Section 9.1.2)** *The number of pinkening steps performed in time interval $(T, 2T)$ is* at least *the number of $b \in V$ that satisfy:*

1. *$b \in I(W)$ (ie. $b$ is white at time $T$);*
2. *there exists some $a \in V$ with $a \in I(R)$ (ie. $a$ is red at time $T$) and $F_a = b$.*

The claim implies:

$$\begin{aligned} |R^C_{2T_-}| &= |R| - \#\text{ of pinkening steps in } (T, 2T] && (9.1.2)\\ &\leq |R| - \sum_{b\in I(W)} \mathbb{I}_{\{\exists a\in I(R)\,:\,F_a=b\}}. && (9.1.3)\end{aligned}$$

We note that for any integer $k \in \mathbb{N}$,

$$\mathbb{I}_{\{k\geq 1\}} \geq k - \binom{k}{2},$$

hence:

$$\begin{aligned} \sum_{b\in I(W)} \mathbb{I}_{\{\exists a\in I(R)\,:\,F_a=b\}} &\geq \sum_{b\in I(W)} \left( |\{a\in I(R)\,:\,F_a=b\}| - \binom{|\{a\in I(R)\,:\,F_a=b\}|}{2} \right) \\ &= \left( \sum_{b\in I(W), a\in I(R)} \mathbb{I}_{\{F_a=b\}} \right) && (9.1.4)\\ &\quad - \left( \sum_{b\in I(W), \{a,a'\}\in I(R)} \mathbb{I}_{\{F_a=b\}\cap\{F_{a'}=b\}} \right). \end{aligned}$$

Thus:

$$\begin{aligned} \mathbb{E}\left[|R^C_{2T}| - |R|\right] \leq &- \sum_{(a,b)\in(V)_2} \mathbb{P}\left(a\in I(R), b\in I(W), F_a=b\right) \\ &+ \sum_{\{a,a'\}\subset V, b\in V} \mathbb{P}\left(a, a'\in I(R), b\in I(W), F_a=b, F_{a'}=b\right). \quad (9.1.5)\end{aligned}$$

The event $\{F_a = b\}$ is entirely determined by the points of the marked Poisson process and by the coin flips performed in the time interval $(T, 2T)$, and therefore is is independent of $I$ (cf. Proposition 9.1.2). We deduce:

$$\begin{aligned} &\sum_{(a,b)\in(V)_2} \mathbb{P}\left(a\in I(R), b\in I(W), F_a=b\right) \\ &\qquad = \sum_{(a,b)\in(V)_2} \mathbb{P}\left(a\in I(R), b\in I(W)\right)\mathbb{P}\left(F_a=b\right) \\ &\text{(use Proposition 9.1.1)} \geq \frac{|R|}{13|V|} \sum_{(a,b)\in(V)_2} \mathbb{P}\left(F_a=b\right) = \frac{|R|}{13|V|} \sum_{a\in V} \mathbb{P}\left(F_a\neq *\right). \quad (9.1.6)\end{aligned}$$

For a given $a \in V$, $\mathbb{P}(F_a = *)$ is the probability that there is no $T < \tau_n < 2T$ with $e_n \ni a$. Notice that this is at most the probability that $\tilde{I}_{2T}(a) = a$: $a$ cannot move if there is no edge $e_n \ni a$ with $T < \tau_n \leq 2T$. We deduce:

$$\mathbb{P}(F_a \neq *) \geq 1 - \mathbb{P}\left(\tilde{I}_{2T}(a) = a\right) = \mathbb{P}\left(a_T^R \neq a\right)$$

where $\{a_t^R\}_{t \geq 0}$ is a realization of $\mathsf{RW}(G)$ started from $a$. By the contraction principle and Proposition 2.3.1,

$$T = 20\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \geq 20\mathsf{T}_{\mathsf{RW}(G)}(1/4) \geq 2^{-20}$$

which implies:

$$\mathbb{P}\left(a_T^R \neq a\right) \geq 1 - \frac{1}{|V|} - 2^{-20} \geq \frac{13}{14} \text{ since } |V| \geq 300.$$

We deduce from (9.1.6) that:

$$\sum_{(a,b) \in V^2} \mathbb{P}(a \in I(R), b \in I(W), F_a = b) \geq \frac{|R|}{14}. \tag{9.1.7}$$

We now consider the second sum in the RHS of (9.1.5). As before, we notice that $\{F_a = b, F_{a'} = b\}$ is independent of $I$ and therefore:

$$\begin{aligned}
&\sum_{\{a,a'\} \subset V, b \in V \setminus \{a,a'\}} \mathbb{P}\left(a, a' \in I(R), b \in I(W), F_a = b, F_{a'} = b\right) \\
&\quad = \sum_{\{a,a'\} \subset V, b \in V \setminus \{a,a'\}} \mathbb{P}\left(a, a' \in I(R), b \in I(W)\right) \mathbb{P}(F_a = b, F_{a'} = b) \\
&\quad \leq \sum_{\{a,a'\} \subset V, b \in V \setminus \{a,a'\}} \mathbb{P}\left(a, a' \in I(R)\right) \mathbb{P}(F_a = b, F_{a'} = b).
\end{aligned}$$

We *claim* that:

**Claim 3 (proven in Section 9.1.3)** *For all $a, a' \in V$ with $a \neq a'$,*

$$\sum_{b \in V \setminus \{a,a'\}} \mathbb{P}(F_a = b, F_{a'} = b) \leq \mathbb{P}\left(F_a = a'\right) + \mathbb{P}(F_{a'} = a).$$

The Claim implies:

$$\begin{aligned}
&\sum_{\{a,a'\}\subset V, b\in V\setminus\{a,a'\}} \mathbb{P}\left(a,a'\in I(R)\right)\mathbb{P}\left(F_a=b,F_{a'}=b\right)\\
&\qquad\leq \sum_{\{a,a'\}\subset V}\mathbb{P}\left(a,a'\in I(R)\right)\left(\mathbb{P}\left(F_a=a'\right)+\mathbb{P}\left(F_{a'}=a\right)\right)\\
&\qquad= \sum_{(a,a')\in(V)_2}\mathbb{P}\left(a\in I(R),a'\in I(R)\right)\mathbb{P}\left(F_a=a'\right)\\
&\text{(apply Prop. 9.1.1)} = \left(\frac{|R|}{|V|}+2^{-10}\right)\sum_{(a,a')\in(V)_2}\mathbb{P}\left(a\in I(R)\right)\mathbb{P}\left(F_a=a'\right)\\
&(\cup_{a'}\{F_a=a'\}=\{F_a\neq *\}) = \left(\frac{|R|}{|V|}+2^{-10}\right)\sum_{a\in V}\mathbb{P}\left(a\in I(R)\right)\mathbb{P}\left(F_a\neq *\right)\\
&(\mathbb{P}\left(F_a\neq *\right)\leq 1) \leq \left(\frac{|R|}{|V|}+2^{-10}\right)\sum_{a\in V}\mathbb{P}\left(a\in I(R)\right)\\
&\qquad= \left(\frac{|R|}{|V|}+2^{-10}\right)\mathbb{E}\left[|I^{-1}(R)|\right]\\
&\qquad= \left(\frac{|R|}{|V|}+2^{-10}\right)|R| \text{ since } I=I_{(0,T]} \text{ is a bijection.}
\end{aligned}$$

Plugging this equation and (9.1.7) into (9.1.5) we obtain:

$$\mathbb{E}\left[|R_{2T}^C|-|R|\right] \leq |R|\left(\frac{|R|}{|V|}+2^{-10}-\frac{1}{14}\right) \leq -|R|/30, \textbf{ if } |R| \leq |V|/28. \quad (9.1.8)$$

If $|R|>|V|/28$, we can still find a subset $R_0\subset R$ of size $|R_0|=\lfloor |V|/28\rfloor$. Going back to (9.1.3) and (9.1.4), we see that:

$$\begin{aligned}
\sum_{b\in I(W)}\mathbb{I}_{\{\exists a\in I(R):F_a=b\}} &\geq \sum_{b\in I(W)}\mathbb{I}_{\{\exists a\in I(R_0):F_a=b\}}\\
&\geq \sum_{b\in I(W)}|\{a\in I(R_0)\,:\,F_a=b\}| - \binom{|\{a\in I(R_0)\,:\,F_a=b\}|}{2}\\
&= \left(\sum_{b\in I(W),a\in I(R_0)}\mathbb{I}_{\{F_a=b\}}\right) \qquad (9.1.9)\\
&\quad -\left(\sum_{b\in I(W),\{a,a'\}\in I(R_0)}\mathbb{I}_{\{F_a=b\}\cap\{F_{a'}=b\}}\right).
\end{aligned}$$

We may repeat the reasoning presented from (9.1.4) onwards, replacing $R$ by $R_0$, to deduce that:

$$\mathbb{E}\left[|R^C_{2T_-}| - |R|\right] \leq -\frac{|R_0|}{30}.$$

We now note that, since $|V| \geq 300$,

$$|R_0| \geq \frac{|V|}{30} - 1 \geq \frac{3|V|}{100} \geq \frac{3|R|}{100}$$

since $|R| \leq |V|$. We deduce that:

$$\mathbb{E}\left[|R^C_{2T_-}| - |R|\right] \leq -\frac{|R|}{1000} \text{ if } |R| > |V|/28,$$

which gives the Lemma together with (9.1.8). □

### 9.1.1 Proof of the required estimates for the $I$ map (Proposition 9.1.1)

*Proof:* [of Proposition 9.1.1]Recall that $T \geq 20\mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$, therefore $T \geq 2\mathsf{T}_{\mathsf{IP}(2,G)}(2^{-10})$ by Proposition 2.3.1. By the contraction principle [2], this also implies that $T \geq \mathsf{T}_{\mathsf{RW}(G)}(2^{-10})$.

Recall that $I = I_{(0,T]}$ as in the construction of the modified chameleon process. This implies that for any set $S$, $I(S)$ has the law of $\mathsf{EX}(|S|, G)$ started from $S$. We deduce:

$$\begin{aligned}
\mathbb{P}\left((a,b) \in I(S) \times I(S)\right) &= \mathbb{P}\left(\{a,b\} \subset S^I_T\right) \\
\text{(negative correlation, Lemma 3.3.1)} &\leq \mathbb{P}\left(a \in S^I_T\right)\mathbb{P}\left(b \in S^I_T\right) \\
(\mathcal{L}[I] = \mathcal{L}[I^{-1}], \text{ Proposition 3.2.2}) &= \mathbb{P}\left(a \in I(S)\right)\mathbb{P}\left(b^I_T \in S\right) \\
(T \geq \mathsf{T}_{\mathsf{RW}(G)}(2^{-10})) &= \mathbb{P}\left(a \in I(S)\right)\left(\frac{|S|}{|V|} + 2^{-10}\right).
\end{aligned}$$

As for the other inequality in the proposition, we have:

$$\begin{aligned}
\mathbb{P}\left((a,b) \in I(S) \times I(L)\right) &= \mathbb{P}\left((a^I_t, b^I_t) \in S \times L\right) \\
(\text{take } \mathbf{x} = (a,b)) &= \mathbb{P}\left(\mathbf{x}^I_t \in S \times L\right) \\
(*) &\geq \left(1 - 2^{-9}\right)^2 \frac{|S \times L|}{|(V)_2|} \\
&\geq \left(1 - 2^{-8}\right) \frac{|S|\,|L|}{|V|^2},
\end{aligned}$$

where $(*)$ follows from the symmetry of the transition rates of $\mathsf{IP}(2,G)$, the fact that $T \geq 2\mathsf{T}_{\mathsf{IP}(2,G)}(2^{-10})$ and Proposition 2.3.2. We note that $|L|/|V| \geq (1/12)$ and $1 - 2^{-8} \geq 12/13$ to finish the proof. □

### 9.1.2 Proof of claim on the number of pinkenings (Claim 2)

*Proof:* [of Claim 2] For each $b$ satisfying the two items, choose the $a = a_b$ with $a \in I(R)$, $F_a = b$ and the *least* value of $\phi_a$ among all such vertices. We will argue that each such red/white pair $(a_b, b)$ corresponds to one pinkening step in $(T, 2T)$.

To prove this, fix $(a_b, b)$ as above. Notice that, since $F_{a_b} = b \neq *$, there exists a $m$ such that $\phi_{a_b} = \tau_m$. Moreover, we have $e_m \equiv \{a_b, \tilde{I}_{\tau_m -}(b)\}$.

We now check that time $\tau_m$ satisfies the conditions for a pinkening time, with the "suitable modification" proposed in the proof of Lemma 9.0.1 (cf. Box 9.1).

- *The process is in a color-changing phase.* Indeed, $\tau_m \in (T, 2T)$.
- *Vertex $a_b$ is red at time $T$ and there is no time $\tau_j \in [T, \tau_m)$ where $a_b \in e_j$.* This follows from the definition of $\phi_{a_b}$, which equals $\tau_m$.
- *The other endpoint of $e_m$ at time $\tau_{m-}$, which is $\tilde{I}_{\tau_m -}(b)$, is white.* For suppose that was not true. Since $\tilde{I}_T(b) = b \in I(W)$ – ie. $b$ was white at time $T$ –, there must have been some time $\tau_j < \tau_m$ at which $\tilde{I}_{\tau_j -}(b)$ was involved in a pinkening step. But then $e_j = \{u, \tilde{I}_{\tau_j -}(b)\}$ for some $u \in I(R)$, and moreover $F_u = b$, since otherwise condition 3 would prevent a pinkening step from happening at time $\tau_j$. But then $u \in I(R)$, $F_u = b$ and $\phi_u = \tau_j < \tau_m = \phi_{a_b}$, in contradiction with the choice of $a_b$.

This shows that each pair $(a_b, b)$ as above will give a pinkening step. □

### 9.1.3 Proof of claim on $F_a$ (Claim 3)

*Proof:* [of Claim 3] We will first show that, for $(a, a', b) \in (V)_3$,

$$\textbf{(Goal \#1)}\ \mathbb{P}\left(F_a = b, F_{a'} = b, \phi_a \leq \phi_{a'}\right) = \mathbb{P}\left(F_a = b, F_{a'} = a, \phi_a \leq \phi_{a'}\right). \tag{9.1.10}$$

Let $L_b, R_b$ denote the events appearing in the LHS and RHS of (9.1.10) (respectively). We present a simple measure-preserving mapping $\Phi$, which acts on

$$(\mathcal{P}, \{e_n\}_n, \{c_n\}_n, \{d_k\}_k),$$

that maps $L_b$ into $R_b$ and vice-versa. We describe $\Phi$ in words: all values of $d_k$, $T < \tau_j \leq 2T$ and all corresponding $e_j$ and $c_j$, *except* for the following modification: if $\tau_m = \phi_a$, we flip the value of $c_m$ to $c'_m = 1 - c_m$.

Let us check that $\Phi$ has the desired properties. $\Phi$ is clearly measure-preserving, since $\phi_a$ is a stopping time that is independent of the value $c_m$ of the flipped coin.

Now suppose $\{\hat{I}_t\}_{T<t\leq 2T}$ is defined precisely as $\{\tilde{I}_t\}_{T<t\leq 2T}$, but with $c_m$ is flipped. The two processes coincide for any time $T < t < \phi_a$. If $L_b$ holds, at time $\phi_a < 2T$, the endpoints

of $e_j$ are $a$ and $\tilde{I}_{\phi_a-}(F_a) = \tilde{I}_{\phi_a-}$ (by definition of $F_a$). Since the coin flips used for $\tilde{I}_{\phi_a}$ and $\hat{I}_{\phi_a}$ are opposite, we have

$$(\hat{I}_{\phi_a}(a), \hat{I}_{\phi_a}(b)) = (\tilde{I}_{\phi_a}(b), \tilde{I}_{\phi_a}(a))$$

whereas $\tilde{I}_{\phi_a}(c) = \hat{I}_{\phi_a}(c)$ for all $c \in V\backslash\{a, F_a\}$. Since the two processes evolve in the same way after time $\phi_a$, we have the following:

$$\hat{I}_t(c) = \begin{cases} \tilde{I}_t(a), & \text{if } c = F_a \text{ and } t \geq \phi_a; \\ \tilde{I}_t(F_a), & \text{if } c = a \text{ and } t \geq \phi_a; \\ \tilde{I}_t(c) & \text{in all other cases.} \end{cases}$$

It is easy to see that $\phi_a, \phi_{a'}$ retain their values and that the random variable $\hat{F}_a$ corresponding to $F_a$ in the $\hat{I}$ process satisfies $\hat{F}_a = F_a$. In the $\tilde{I}_t$ process, $F_{a'} = b$ meant that the edge $e_k$ correspondng to $\tau_k = \phi_{a'}$ was of the form $e_k = \{a', \tilde{I}_{\phi_{a'}-}(b)\}$; but then, in the $\hat{I}$ process we have $e_k = \{a', \hat{I}_{\phi_{a'}-}(a)\}$ since $F_a = b$. That is, the value of $F_{a'}$ in the $\hat{I}_t$ process is $a$. We deduce that the $\hat{I}$ process lies in the event $R_b = \{F_a = b, F_{a'} = a, \phi_a \leq \phi_{a'}\}$. Since $\Phi$ is measure preserving, this shows that: $\mathbb{P}(L_b) = \mathbb{P}(R_b)$, ie. (9.1.10) holds.

Now sum over $b$ the LHS and the RHS of (9.1.10):

$$\sum_{b \in V\backslash\{a,a'\}} \mathbb{P}(F_a = b, F_{a'} = b, \phi_a \leq \phi_{a'})$$
$$\leq \sum_{b \in V\backslash\{a,a'\}} \mathbb{P}(F_a = b, F_{a'} = a, \phi_a \leq \phi_{a'}) \leq \mathbb{P}(F_{a'} = a). \quad (9.1.11)$$

Similarly,

$$\sum_{b \in V\backslash\{a,a'\}} \mathbb{P}(F_a = b, F_{a'} = b, \phi_a > \phi_{a'}) \leq \mathbb{P}\left(F_a = a'\right). \quad (9.1.12)$$

Adding up (9.1.11) and (9.1.12) finishes the proof. □

**Remark 9.1.2** *Notice that this proof only works because a red vertex will only become pink if the* first *edge it ever touches also touches a white particle. This is what makes the definition of $F_a$ useful, and only with this definition does the previous proof work.*

## 9.2 Estimate for the first depinking time (Lemma 9.0.2)

*Proof:* [of Lemma 9.0.2] We will work with the "suitably modified process" from the proof of Lemma 9.0.1, ie. the rules for a PINKENING TIME are as in Box 9.1. Notice that this change does *not* change the value (or the distribution) of $D_1$, since:

$$D_1 \equiv \inf\{2kT \,:\, k \in \mathbb{N}\backslash\{0\}, |P^C_{2kT-}| \geq \min\{|R^C_{2kT-}|, |W^C_{2kT-}|\}\}$$

and this condition continues to hold if one allows pinkening steps even after $|P^C_{2kT_-}|$ is at least as large as the minimum. The modification to the process also does not affect the end result of Lemma 5.3.1: that is, the discrete-time process starting from $(\mathbf{z}, R, P, W)$ with subsequent states $(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i)$ described in that Lemma is a time-homogeneous Markov chain, and $\hat{D}_1 \equiv D_1/2T$ is a stopping time for this process.

We will also assume that $|R| \leq |W|$; this is without loss of generality, since the roles of white and red particles in the chameleon process is symmetric. Since $P$ is empty and $\mathbf{O}(\mathbf{z}) \cup R \cup P \cup W = V$, it follows that: $|W| \geq (|V| - k)/2 \geq |V|/4$ since $k \leq |V|/2$. At any time $2kT < D_1$, the number of red and white particles are $|R| - p_k$ and $|W| - p_k$, where $p_k < \Delta(|R|)$ (recall Proposition 7.1.1 for the definition of $\Delta$); otherwise, the would have been a depinking at the *first time* $2jT$ with $j \leq k$ and $p_j \geq \Delta(|R|)$ (cf. the proof of Proposition 7.1.1). This discussion implies that the number of white particles at times $2kT < D_1$ is at least $|W| - \Delta(|R|) + 1 \geq |W| - |R|/3 \geq 2|W|/3 \geq |V|/12$. In particular, this implies:

$$\forall k < \hat{D}_1, : |\hat{W}_k| \geq |V|/12. \tag{9.2.1}$$

Notice that $\hat{D}_1 > k \Rightarrow |\hat{R}_k| > |R|/3$. In other words, if $|R^C_{2jT_-}| \leq |R|/3$ for some $j \leq k$, that would mean that the number of pink particles would exceed that of red particles at time $2jT$, which would lead to a depinking at that time. We deduce:

$$\begin{aligned}
\forall k \in \mathbb{N}\backslash\{0\},\ \mathbb{P}\left(\hat{D}_1 > k\right) &\leq \frac{3\mathbb{E}\left[\hat{R}_k \mathbb{I}_{\{\hat{D}_1 > k\}}\right]}{|R|} \\
&\leq \frac{3\mathbb{E}\left[\hat{R}_k \mathbb{I}_{\{\hat{D}_1 > (k-1)\}}\right]}{|R|} \\
&= \frac{3\mathbb{E}\left[\mathbb{E}\left[\hat{R}_k \mid \hat{\mathcal{F}}_{k-1}\right] \mathbb{I}_{\{\hat{D}_1 > (k-1)\}}\right]}{|R|}.
\end{aligned} \tag{9.2.2}$$

where $\hat{\mathcal{F}}_{k-1}$ is the $\sigma$-field generated by $(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i)$ for $i \leq k-1$.

We now estimate the integrand in (9.2.2). Lemma 5.3.1 and its proof implies that $\mathbb{E}\left[\hat{R}_k \mid \hat{\mathcal{F}}_{k-1}\right]$ is the expected number of red particles after a potential depinking, a constant-color phase and a color-changing phase for a chameleon process started from

$$(\hat{\mathbf{z}}_{k-1}, \hat{R}_{k-1}, \hat{P}_{k-1}, \hat{W}_{k-1}) \in \mathcal{C}_k(V).$$

By Lemma 9.0.1, we can ensure that:

$$\mathbb{E}\left[\hat{R}_k \mid \hat{\mathcal{F}}_{k-1}\right] \leq (1-c)\hat{R}_{k-1} \text{ if } |\hat{W}_{k-1}| \geq |V|/12 \text{ and } |P^C_{2(k-1)T}| < |R^C_{2(k-1)T}|.$$

These conditions are always satisfied in the event $\{\hat{D} > (k-1)\}$, by (9.2.1), and outside this event the integrand of (9.2.2) is 0. We deduce from all this that:

$$\begin{aligned}
\forall k \in \mathbb{N}\backslash\{0\},\ \frac{3\mathbb{E}\left[\hat{R}_k \mathbb{I}_{\{\hat{D}_1 > k\}}\right]}{|R|} &\leq \frac{3\mathbb{E}\left[\mathbb{E}\left[\hat{R}_k \mid \hat{\mathcal{F}}_{k-1}\right] \mathbb{I}_{\{\hat{D}_1 > (k-1)\}}\right]}{|R|} \\
&\leq (1-c)\left\{\frac{3\mathbb{E}\left[\hat{R}_{k-1}\mathbb{I}_{\{\hat{D}_1 > k-1\}}\right]}{|R|}\right\} \\
(...\text{induction}...) &\leq 3(1-c)^k.
\end{aligned}$$

This implies:

$$\mathbb{P}\left(D_1 > 2kT\right) = \mathbb{P}\left(\hat{D}_1 > k\right) \leq 3(1-c)^k,\ c = 1/1000 \text{ universal.}$$

From this one can easily show that $\mathbb{E}\left[e^{D_1/KT}\right] \leq e$ for some universal $K$. $\square$

## 9.3 Proof of Lemma 6.2.2

*Proof:* Fix $\mathbf{x} \in (V)_k$. We first prove that:

$$\mathbb{E}\left[e^{\frac{D_j(\mathbf{x})}{K_2 T}}\right] \leq e^j,\ K_2 > 0 \text{ universal.} \tag{9.3.1}$$

this is the bound we wish to obtain except that we are not conditioning on Fill.

We proceed as in the previous proof and consider the discrete-time process

$$\{(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i)\}_{i \geq 0}$$

introduced in Lemma 5.3.1, henceforth called the hat process. This time we take the initial state

$$(\mathbf{z}, R, P, W) \equiv (\mathbf{z}, \{x\}, \emptyset, V\backslash(\mathbf{O}(\mathbf{z}) \cup \{x\}))$$

corresponding to $\mathbf{x} = (\mathbf{z}, x)$ in the sense of Proposition 5.4.1. Also recall the definition $\hat{D}_i \equiv D_i(\mathbf{x})/2T$ and note that (9.3.1) is equivalent to:

$$\mathbb{E}\left[e^{\frac{\hat{D}_j}{K_3 T}}\right] \leq e^j,\ K_3 > 0 \text{ universal.} \tag{9.3.2}$$

This is valid for $j = 1$ due to Lemma 9.0.2. For $j > 1$, we recall the definition of the $\sigma$-fields $\hat{\mathcal{F}}_i$, recall that $\hat{D}_{j-1}$ is a stopping time for the hat process (cf. Lemma 5.3.1) and obtain:

$$\mathbb{E}\left[e^{\frac{\hat{D}_j}{K_3 T}}\right] \leq \mathbb{E}\left[e^{\frac{\hat{D}_{j-1}}{K_3 T}}\,\mathbb{E}\left[e^{\frac{\hat{D}_j - \hat{D}_{j-1}}{K_3 T}} \mid \hat{\mathcal{F}}_{\hat{D}_{j-1}}\right]\right]. \tag{9.3.3}$$

We will apply the strong Markov property of the hat process (cf. Lemma 5.3.1 again) to bound the conditional expectation in the RHS. Indeed, we note that the conditional law of the process given $\hat{\mathcal{F}}_{\hat{D}_{j-1}}$ is the law of the hat process started from state:

$$(\hat{\mathbf{z}}_{\hat{D}_{j-1}}, \hat{R}_{\hat{D}_{j-1}}, \hat{P}_{\hat{D}_{j-1}}, \hat{W}_{\hat{D}_{j-1}}).$$

Notice that $\hat{P}_{\hat{D}_{j-1}} = P^C_{D_{j-1}\_} \neq \emptyset$; in fact, since depinking occurs at time $D_{j-1}$, we know that $|P^C_{D_{j-1}\_}| \geq \min\{|R^C_{D_{j-1}\_}|, |W^C_{D_{j-1}\_}|\}$. However, at time $D_{j-1}$ all pink particles disappear: the hat process evolves as if started from a state with no pink particles, and $\hat{D}_j - \hat{D}_{j-1}$ is the first depinking time for the hat process with this modified initial state. We deduce from Lemma 9.0.2 that:

$$\mathbb{E}\left[ e^{\frac{\hat{D}_j - \hat{D}_{j-1}}{K_3 T}} \mid \hat{\mathcal{F}}_{\hat{D}_{j-1}} \right] \leq e \text{ almost surely,}$$

so that:

$$\mathbb{E}\left[ e^{\frac{\hat{D}_j}{K_3 T}} \right] \leq \mathbb{E}\left[ e^{\frac{\hat{D}_{j-1}}{K_3 T}} \right] e \leq e^j \text{ by induction.} \tag{9.3.4}$$

This proves (9.3.3) and (9.3.1).

To prove the Lemma, notice that conditioning on $\mathsf{Fill}$ simply biases the coin flips $d_k$ performed at depinking times (cf. Remark 7.2.1). This will not change the distribution of $\hat{D}_1$ or the conditional distribution of $\hat{D}_j - \hat{D}_{j-1}$ given the past of the process, so the argument we presented above still applies. □

# 10 Applications and comparison with previous results

In this section we present some applications of the main theorem and show how they represent improvements over the previous best bounds. For simplicity, we will mostly consider the cases of symmetric exclusion where the number of particles $k$ equals $c|V|$, for some constant $c > 0$; the asymptotics might be slightly different in other cases. We will also deal exclusively with the 1/4-mixing time.

We will use asymptotic notation for shorthand: whenever we write $a = O(b)$ (or $b = \Omega(a)$), we simply mean that $a \leq C\,b$ for some universal $C > 0$; whereas $a = \Theta(b)$ means that $b/c \leq a \leq cb$ for some universal $c > 0$.

## 10.1 Setting up a fair comparison with previous results

Let $G = (V, E, \{w_e\}_{e \in E})$ be a general weighted graph with $n = |V|$ vertices. It seems that the best *general* bound that was previously available (implicitly) for the mixing time of $\mathsf{EX}(k, G)$ comes from the combination of three ingredients.

**Mixing time from Log-Sobolev constant.** Recall that the state space of $\mathsf{EX}(k,G)$ has cardinality $\binom{n}{k} = 2^{\Theta(n)}$ if $k = \Theta(n)$. By the results of [10], if $\rho_{\mathsf{EX}(k,G)}$ is the log-Sobolev constant of $\mathsf{EX}(k,G)$, then:

$$\mathsf{T}_{\mathsf{EX}(k,G)}(1/4) = O\left(\frac{\ln n}{\rho_{\mathsf{EX}(k,G)}}\right) \text{ for } k = \Theta(n).$$

**Log-Sobolev inequality for the Bernoulli-Laplace model.** Consider the complete graph $K_n$ on $n$ vertices where each edge has weight $1/n$. We call the corresponding weighted graph $\mathcal{K}_n$. $\mathsf{EX}(k,\mathcal{K}_n)$ is the so-called *Bernoulli-Laplace model* with $k$ particles. Lee and Yau [12, Theorem 5] showed that the log-Sobolev constant of $\mathsf{EX}(k,\mathcal{K}_n)$ is $\Theta\left(\ln(n^2/k(n-k))\right)$. Notice that this is $\Theta(1)$ for $k = \Theta(n)$

**Comparison argument.** Now consider a general weighted graph $G = (V, E, \{w_e\}_{e\in E})$ with $n = |V|$ vertices. Assume that for each pair $(x,y) \in V^2$ one has defined a path $\gamma_{x,y}$ in $G$ connecting $x$ to $y$. For each such pair, let $I_{x,y}(e) = 1$ if $e$ is crossed by $\gamma_{x,y}$ and 0 otherwise, and also let $\ell_{x,y}$ denote the length of $\gamma_{x,y}$. Finally, define:

$$\phi(G) \equiv \max_{e\in E} \sum_{(x,y)\in V^2} \frac{1}{nw_e} I_{x,y}(e).$$

The *comparison* argument of Diaconis and Saloff Coste [9, 10] implies that the log-Sobolev of $\mathsf{RW}(G)$ is $\Omega(1/\phi(G))$ times that of $\mathsf{RW}(\mathcal{K}_n)$. A non-trivial extension of this result, also presented in [9], shows that the same result holds for the log-Sobolev constants of the corresponding exclusion processes.

We deduce that:

$$\mathsf{T}_{\mathsf{EX}(k,G)}(1/4) = O\left(\phi(G)\ \ln|V|\right) \text{ if } k = \Theta(|V|), \text{ where } \phi(G) \text{ is as defined above.} \quad (10.1.1)$$

It can be very hard to find good upper bounds on $\phi(G)$ in general, but we note the following simple lower bound:

$$\phi(G) \geq \frac{2\overline{\mathrm{dist}^2}}{\overline{d}}, \quad (10.1.2)$$

where $\overline{\mathrm{dist}^2}$ is the average over all $(x,y)\nu V^2$ of the square of the graph-theoretic distance between $x$ and $y$, and $\overline{d}$ is the average (weighted) degree in $G$. Indeed, it suffices to see

that:

$$\begin{aligned}
\phi(G) &\geq \sum_{e\in E} \frac{w_e}{\sum_{f\in E} w_f} \left( \sum_{(x,y)\in V^2} \frac{1}{nw_e} I_{x,y}(e) \right) \\
&= \frac{1}{\sum_{f\in E} w_f} \left[ \sum_{(x,y)\in V^2} \left( \frac{\sum_{e\in E} I_{x,y}(e)\ell_{x,y}}{n} \right) \right] \\
(\text{use } \textstyle\sum_e I_{x,y(e)} = \ell_{x,y}) &= \frac{1}{n^{-1}\sum_{f\in E} w_f} \left[ \sum_{(x,y)\in V^2} \frac{\ell_{x,y}^2}{n^2} \right] \\
(\text{use } \ell_{x,y} \geq \mathrm{dist}(x,y)) &\geq \frac{1}{n^{-1}\sum_{f\in E} w_f} \left[ \sum_{(x,y)\in V^2} \frac{\mathrm{dist(x,y)}^2}{n^2} \right] = \frac{2\overline{\mathrm{dist}^2}}{\overline{d}}.
\end{aligned}$$

## 10.2 Morris' result for $(\mathbb{Z}/L\mathbb{Z})^d$ and $\{1,2,3,\ldots,L\}^d$

Here we consider $G = (\mathbb{Z}/L\mathbb{Z})^d$ or $\{1,2,3,\ldots,L\}^d$ with the usual nearest-neighbor bonds and all edge weights equal to 1. The optimal result in this setting is due to Morris [20], who proved a $O\left(L^2 \ln k\right)$ bound $\mathsf{T}_{\mathsf{EX}(k,G)}(1/4)$ ($d$ fixed, $L, k$ potentially large). Our bound here is $O\left(L^2 \ln L\right)$, which is worse for $\ln k \ll \ln L$, but matches the previous best, which could be deduced from the results in Section 10.1. See also [23] for a much more general result.

## 10.3 Percolation clusters

We now consider a generalization of the above case where Morris' result does not apply. Perform *bond percolation* over $G = \{1,\ldots,L\}^d$; that is, remove each edge in this graph independently with probability $1-p$, where $p$ is greater than the critical parameter $p_c(\mathbb{Z}^d)$ for bond percolation in $\mathbb{Z}^d$. It is known that the resulting random graph typically has one single connected component $C_{L,d,p}$ of size $\Theta\left(L^d\right)$; all other components have sublinear size. The mixing time of $\mathsf{RW}(C_{L,d,p})$ was studied in [5, 22], where it was shown that:

$$\mathsf{T}_{\mathsf{RW}(C_{L,d,p})}(1/4) = \Theta\left(L^2\right) \text{ with probability tending to 1 as } L \to +\infty.$$

Theorem 1.1.1 immediately gives:

$$\text{“}\forall k,\ \forall \epsilon\ :\ \mathsf{T}_{\mathsf{EX}(k,C_{L,d,p})}(\epsilon) \leq O\left(L^2 \ln(L/\epsilon)\right)\text{” with probability tending to 1 as } L \to +\infty.$$

It does not seem possible to reobtain this result via the techniques in Section 10.1.

## 10.4 Bounded-degree expanders

Consider the family of all weighted graphs $G = (V, E, \{w_e\})$ with bounded degree and where the spectral gap of random walk is bounded below by a universal $\delta > 0$. We call these graphs *expanders*, in a slight extension of normal usage. It is well-known that random walk on such graphs has mixing time $\Theta(\ln |V|)$, so our result implies:

$$\forall G \text{ as above, } \mathsf{T}_{\mathsf{EX}(c|V|,G)}(1/4) \leq C_\delta \ln(|V|), \text{ where } C_\delta \text{ depends only on } \delta.$$

By contrast, one easily sees that $\overline{\text{dist}}^2 = \Omega\left(\ln^2 |V|\right)$, so (10.1.2) implies that the best mixing time bound that one could derive from (10.1.1) is of the order $O\left(\ln^3 |V|\right)$ for $k = c|V|$. In fact it is not clear whether this can be achieved in general, since this requires a suitable construction of paths in $G$.

## 10.5 The giant component of the Erdös-Rényi graph

Our next example is the Erdös-Rényi random graph model. This is the random graph $G_{n,p}$ with vertex set $[n]$ and (random) edge set $E_{n,p}$, where each potential edge $\{i, j\} \in \binom{[n]}{2}$ appears with probability $p$, independently from all other potential edges.

We will focus on the case where $p = p(n) = c/n$ where $c > 1$ is a constant and $n \to +\infty$. It is known that the largest connected component $C_1 = C_1(G_{n,p})$ has $\Theta(n)$ vertices and all other components have size $O(\ln n)$ with probability tending to 1 as $n \to +\infty$. The component $C_1$ is known as the *giant component* of $G_{n,c/n}$ and we will consider it as a weighted graph where all edges have weight 1.

The mixing time of the standard random walk in $C_1$ – ie. the process which moves in discrete time from a vertex to a uniformly chosen neighbor – is known to be $O\left(\ln^2 n\right)$ with high probability for $n$ large [4]. It is not hard to show via the relationship between mixing and hitting times derived by Aldous [1] that the continuous-time random walk $\mathsf{RW}(C_1)$ also has mixing time $O\left(\ln^2 n\right)$ with high probability. It follows from Theorem 1.1.1 that:

$$T_{\mathsf{EX}(C_1)}(1/4) = O\left(\ln^3 n\right) \text{ with high probability as } n \to +\infty.$$

A simple calculation shows that the lower bound for $\phi(G)$ that can be obtained from (10.1.2) in this case is $\Omega\left(\ln^2 n\right)$, so $\Omega\left(\ln^3 n\right)$ is the lowest mixing time bound we could possibly achieve via (10.1.1). In this sense, Theorem 1.1.1 is at least as good as the comparison bound in this example. We emphasize that it is not at all clear that one may choose paths over $C_1$ in order to obtain a $O\left(\ln^3 n\right)$ bound for the mixing time.

## 10.6 Graphs on point processes

Our final family of examples consists of random weighted graphs $\mathcal{G}_{L,d,\alpha}$ whose mixing time was determined by Caputo and Faggionato [6]. Fix some $\alpha > 0$ and $d \in \mathbb{N}$. Let $\mathcal{P}$ denote

a Poisson process on $\mathbb{R}^d$. The vertex set $V(\mathcal{G}_{L,d})$ is the set of all points of $\mathcal{P}$ contained in the box $[-L/2, L/2]^d$. The set of edges consists of all pairs of points in $V(\mathcal{G}_{L,d})$, and each edge $\{a, b\}$ has weight equal to

$$w_{\{a,b\}} \equiv e^{-|a-b|^\alpha}.$$

It is easy to see that $|V(\mathcal{G}_{L,d})|/L^d \to 1$ in probability when $L \to +\infty$. It is shown in [6, Theorem 1.2] that the mixing time of $\mathsf{RW}(\mathcal{G}_{L,d,\alpha})$ for $\alpha > d$ satisfies:

$$\exists C = C_{d,\alpha} > 0, \lim_{L\to+\infty} \mathbb{P}\left(T_{\mathsf{RW}(\mathcal{G}_{L,d,\alpha})}(1/4) \le C\, L^2\right) = 1.$$

Theorem 1.1.1 immediately implies:

$$\lim_{L\to+\infty} \mathbb{P}\left(\forall 0 < \epsilon < 1/2,\, \forall 0 \le k \le |V(\mathcal{G}_{L,d,\alpha})|,\, T_{\mathsf{EX}(\mathcal{G}_{L,d,\alpha})}(1/4) \le C_2\, L^2 \ln(L/\epsilon)\right) = 1$$

where $C_2$ depends only on $d$ and $\alpha < d$. Related results can be obtained for $d = \alpha$ if one increases the density of the Poisson process; see [6] for details. As in the previous case, it is not at all clear how alternative techniques could be applied here.

# 11 Final remarks

Our paper leaves many questions open. Here we present a few problems that seem especially interesting:

- Are there any other interacting particle systems whose mixing time can be bounded solely in terms of the constituent parts? Non-symmetric exclusion is an obvious candidate. Another is the zero-range process. Morris [18] used the comparison principle and a coupling argument on the complete graph to bound the spectral gap of this process on a grid. It is possible to bypass the comparison argument and bound mixing directly on an arbitrary graph?

- Can we find a mixing time upper bound of $\mathsf{IP}(|V|, G)$ (i.e. as many particles as vertices) that is similar to our main Theorem? Inspection of the chameleon process shows that it gives the conditional distribution of a particle given the *whole past trajectory* of the other particles. This means, in particular, that it cannot deal with $n$ particles.

- Recall the heuristic assumption in the introduction: $\mathsf{T}_{\mathsf{EX}(k,G)}(\epsilon) \le C_1\, \mathsf{T}_{\mathsf{RW}(G)}(\epsilon/k)$ with $C_1 > 0$ universal. Is this actually true? This would be stronger than our main Theorem.

- Combining the previous two items: is it true that $T_{\mathsf{IP}(|V|,G)}(\epsilon) = C_1\, \mathsf{T}_{\mathsf{RW}(G)}(\epsilon/|V|)$? Could it even be possible that $T_{\mathsf{IP}(|V|,G)}(\epsilon) \leq T_{\mathsf{RW}(|V|,G)}(\epsilon)$, i.e. the interchange process mixes at least as fast as independent random walkers? This would give Aldous' (now proven) conjecture on the spectral gap as a corollary.

# A Some results on ink

## A.1 Proof of Proposition 5.4.1

*Proof:* [of Proposition 5.4.1] We use the modified graphical construction of $\mathbf{x}_t^I$ presented in Section 5.1, which uses the same $\mathcal{P}, \{e_n\}_{n\in\mathbb{N}}$ and $\{c_n\}_{n\in\mathbb{N}}$ that appear in the construction of the chameleon process. We let $\mathcal{G}$ be the $\sigma$-field generated by $\mathcal{P}, \{e_n\}_{n\in\mathbb{N}}$ and prove the following stronger statement:

$$\forall t \geq 0,\, \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\, \mathbb{P}\left(\mathbf{x}_t^I = \mathbf{b} \mid \mathcal{G}\right) = \mathbb{E}\left[\mathsf{ink}_t(b)\mathbb{I}_{\{\mathbf{z}_{\eta_i}^C = \mathbf{c}\}} \mid \mathcal{G}\right]. \tag{A.1.1}$$

Notice that the times $\{\eta_i\}_{i\in\mathbb{N}\cup\{0\}}$ are $\mathcal{G}$-measurable, thus (A.1.1) follows from:

$$\forall i \in \mathbb{N}\cup\{0\},\, \forall \mathbf{b} = (\mathbf{c}, b) \in (V)_k,\, \mathbb{P}\left(\mathbf{x}_{\eta_i}^I = \mathbf{b} \mid \mathcal{G}\right) = \mathbb{E}\left[\rho_{\eta_i}(b)\mathbb{I}_{\{\mathbf{z}_{\eta_i}^C = \mathbf{c}\}} \mid \mathcal{G}\right]. \tag{A.1.2}$$

In fact, we will prove this in tandem with another useful fact:

$$\forall i \in \mathbb{N}\cup\{0\}, \mathbf{z}_{\eta_i}^I = \mathbf{z}_{\eta_i}^C. \tag{A.1.3}$$

Both (A.1.2) and (A.1.3) are clearly true for $i = 0$. Assume inductively that they both hold for $0 \leq i \leq N-1$ for some $N \in \mathbb{N}$. We now consider two possible cases, in what amounts to a somewhat tedious case analysis.

**Case 1: $\eta_N = 2kT$ for some $k \in \mathbb{N}$.** In this case $\mathbf{z}_{\eta_N}^I = \mathbf{z}_{\eta_{N-1}}^I$ and $x_{\eta_N}^I = x_{\eta_{N-1}}^I$, as the $\mathsf{IP}(k, G)$ process does not change at this time. The update rule of the chameleon process implies that $\mathbf{z}_{\eta_N}^C = \mathbf{z}_{\eta_{N-1}}^C$ as well, so that $\mathbf{z}_{\eta_N}^C = \mathbf{z}_{\eta_N}^C$ (i.e. (A.1.3) holds). As for $\mathsf{ink}_{\eta_{N-1}}(v)$, it may or may not change, depending on the color of $v$ at time $\eta_{i-1}$ and on whether or not depinking is performed; but one can easily check that the *expected* change in $\mathsf{ink}_{\eta_{N-1}}(v)$ is 0: each site with 1/2 units of ink changes to one or zero units with equal probability. Thus for $\mathbf{b} = (\mathbf{c}, b) \in (V)_k$,

$$\begin{aligned}
\mathbb{P}\left(\mathbf{x}_{\eta_N}^I = \mathbf{b} \mid \mathcal{G}\right) &= \mathbb{P}\left(\mathbf{z}_{\eta_N}^I = \mathbf{c}, x_{\eta_N}^I = b \mid \mathcal{G}\right) \\
\text{(no change at } \eta_i = \ell T) &= \mathbb{P}\left(\mathbf{z}_{\eta_{N-1}}^I = \mathbf{c}, x_{\eta_{N-1}}^I = b \mid \mathcal{G}\right) \\
\text{(inductive hyp.)} &= \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_{\eta_{N-1}}^C = \mathbf{c}\}}\, \mathsf{ink}_{\eta_{N-1}}(b) \mid \mathcal{G}\right] \\
\text{(zero expected change at } \eta_i = \ell T) &= \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_{\eta_N}^C = \mathbf{c}\}}\, \mathsf{ink}_{\eta_N}(b) \mid \mathcal{G}\right],
\end{aligned}$$

i.e. (A.1.2) holds for $i = N$ in this case.

**Case 2: $\eta_N = \tau_j$ for some $j \in \mathbb{N}$.** Let us introduce the sigma-field $\mathcal{G}_{N-1}$ generated by $\mathcal{G}$, by all $c_n$ with $\tau_N \leq \eta_{N-1}$ and all $d_k$ with $\ell T \leq \eta_{N-1}$. The following claim is not hard to verify (proof omitted).

**Claim 4** *$(\mathbf{z}^C_{\eta_{N-1}}, R^C_{\eta_{N-1}}, P^C_{\eta_{N-1}}, W^C_{\eta_{N-1}})$ are $\mathcal{G}_{N-1}$-measurable, as is the event that $\eta_N \equiv \tau_j$ is a* PINKENING TIME *(cf. Box 5.2). Moreover, $c_j$ is independent of $\mathcal{G}_{N-1}$.*

Our next goal is to prove that:

**Claim 5** *Assuming $\eta_N = \tau_j$ is not a pinkening time, we have the following three equalities:* $\mathbf{z}^C_{\eta_N} = \mathbf{z}^I_{\eta_N}$,

$$\mathbb{P}\left(\mathbf{z}^I_{\eta_N} = \mathbf{c}, x^I_{\eta_N} = x \mid \mathcal{G}_{N-1}\right) = \frac{1}{2}\left(\mathbb{I}_{\{\mathbf{z}^I_{\eta_{N-1}}=\mathbf{c}, x^I_{\eta_{N-1}}=x\}} + \mathbb{I}_{\{\mathbf{z}^I_{\eta_{N-1}}=f_{e_j}(\mathbf{c}), x^I_{\eta_{N-1}}=f_{e_j}(x)\}}\right), \text{ and}$$

$$\mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}^C_{\eta_N}=\mathbf{c}\}}\mathsf{ink}_{\eta_N}(x) \mid \mathcal{G}_{N-1}\right] = \frac{1}{2}\left(\mathbb{I}_{\{\mathbf{z}^C_{\eta_{N-1}}=\mathbf{c}\}}\mathsf{ink}_{\eta_{N-1}}(x) + \mathbb{I}_{\{\mathbf{z}^C_{\eta_{N-1}}=f_{e_j}(\mathbf{c})\}}\mathsf{ink}_{\eta_{N-1}}(f_{e_j}(x))\right).$$

*Proof:* [of Claim 5] $(\mathbf{z}^I_{\eta_N}, x^I_{\eta_N})$ is either $(\mathbf{z}^C_{\eta_{N-1}}, x^I_{\eta_{N-1}})$ or $(f_{e_j}(\mathbf{z}^C_{\eta_{N-1}}), f_{e_j}(x^I_{\eta_{N-1}}))$, depending on whether $c_j = 0$ or 1 (resp.). Moreover, lines 10 – 19 of the pseudocode in Box 5.2 show that the same holds for $\mathbf{z}^I_{\eta_N}$. Since $\mathbf{z}^I_{\eta_{N-1}} = \mathbf{z}^C_{\eta_{N-1}}$ by the induction hypothesis, the first statement in the Claim follows.

The second equality follows from the fact that $c_j$ is independent from $\mathcal{G}_{N-1}$ and therefore the two possibilities for $(\mathbf{z}^I_{\eta_N}, x^I_{\eta_N})$ are equally likely.

As for the third equality, we note that $c_j = 0$ implies $\mathbf{z}^I_{\eta_{N-1}} = \mathbf{z}^I_{\eta_N}$ and $\mathsf{ink}_{\eta_{N-1}} = \mathsf{ink}_{\eta_N}$ (as the sets of red, white and pink particles do not change). For $c_j = 1$, however, $f_{e_j}(\mathbf{z}^I_{\eta_{N-1}}) = \mathbf{z}^I_{\eta_N}$ and $\mathsf{ink}_{\eta_{N-1}} \circ f_{e_j} = \mathsf{ink}_{\eta_N}$, as one can check via from lines 10 - 15 in Box 5.2 and the definition of $\mathsf{ink}_t(v)$ (5.4.2). Thus the same reasoning just presented establishes the third inequality in the Claim. □

We also claim that a similar phenomenon happens at pinkening times.

**Claim 6** *The same conclusions of Claim 5 hold if $\eta_N = \tau_j$ is a pinkening time.*

*Proof:* [of Claim 6] The conditions for pinkening (cf. Box 5.2) show that, at any pinkening time, the edge where pinkening will take place has one red endpoint and one white endpoint. In particular, no endpoint of $e_j$ is black, i.e. $e_j$ does not intersect the coordinates of $\mathbf{z}^I_{\eta_{N-1}}$. That shows that $f_{e_j}(\mathbf{z}^I_{\eta_{N-1}}) = \mathbf{z}^I_{\eta_{N-1}}$ and thus $\mathbf{z}^I_{\eta_N} = \mathbf{z}^I_{\eta_{N-1}}$ regardless of whether $c_j = 0$ or 1. Since it is also the case that $\mathbf{z}^C_{\eta_N} = \mathbf{z}^C_{\eta_{N-1}}$ (cf. lines 4 - 9 in Box 5.2), the inductive hypothesis implies $\mathbf{z}^C_{\eta_N} = \mathbf{z}^I_{\eta_N}$.

The second equality in the claim follows as in the case where there is no pinkening.

Finally, for the third equality, we note that the endpoints of $e_j$ has 0 or 1 units of ink and now both have 1/2 a unit; whereas the amount of ink at other vertices does not change. This corresponding to setting $\mathsf{ink}_{\eta_N}(x) = \mathsf{ink}_{\eta_{N-1}}(f_{e_j}(x))/2 + \mathsf{ink}_{\eta_{N-1}}(x)/2$ for all vertices. Since $\mathbf{z}^C_{\eta_N} = \mathbf{z}^C_{\eta_{N-1}} = f_{e_j}(\mathbf{z}^C_{\eta_{N-1}})$, this gives the desired equality. $\square$

All that is left to do is to put the last two claims together. We see that in either case we can guarantee that (A.1.3) holds at time $i = N$. Moreover, taking expected values conditionally on $\mathcal{G}$ (which is contained in $\mathcal{G}_{N-1}$), we see that:

$$\begin{aligned}
\mathbb{P}\left(\mathbf{z}^I_{\eta_N} = \mathbf{c}, x^I_{\eta_N} = x \mid \mathcal{G}\right) &= \mathbb{E}\left[\mathbb{P}\left(\mathbf{z}^I_{\eta_N} = \mathbf{c}, x^I_{\eta_N} = x \mid \mathcal{G}_{N-1}\right) \mid \mathcal{G}\right] \\
\text{(use two claims)} &= \frac{1}{2}\mathbb{P}\left(\mathbf{z}^I_{\eta_{N-1}} = \mathbf{c}, x^I_{\eta_{N-1}} = x \mid \mathcal{G}\right) \\
&\quad + \mathbb{P}\left(\mathbf{z}^I_{\eta_{N-1}} = f_{e_j}(\mathbf{c}), x^I_{\eta_{N-1}} = f_{e_j}(x) \mid \mathcal{G}\right) \\
\text{(inductive hyp. )} &= \frac{1}{2}\mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}^I_{\eta_{N-1}} = \mathbf{c}\}}\mathsf{ink}_{\eta_{N-1}}(x) \mid \mathcal{G}\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}^I_{\eta_{N-1}} = f_{e_j}(\mathbf{c})\}}\mathsf{ink}_{\eta_{N-1}}(f_{e_j}(x)) \mid \mathcal{G}\right] \\
\text{(two claims again)} &= \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}^I_{\eta_N} = \mathbf{c}\}}\mathsf{ink}_{\eta_N}(x) \mid \mathcal{G}\right].
\end{aligned}$$

[We observe that we have tacitly used the fact that $e_j$ is $\mathcal{G}$-measurable, which comes from the definition of $\mathcal{G}$.] Thus (A.1.2) also holds for $i = N$ and therefore for all $i \in \mathbb{N} \cap \{0\}$, as desired. $\square$

# References


[1] David Aldous. Some inequalities for reversible Markov chains. *Journal of the London Mathematical Society*, 25(3):564–576, 1982.

[2] David Aldous and James Allen Fill. Reversible Markov Chains and Random Walks on Graphs. Book draft, `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[3] Enrique D. Andjel. A correlation inequality for the symmetric exclusion process. *Annals of Probability*, 16(2):717–721, 1988.

[4] Itai Benjamini, Gady Kozma, and Nicholas Wormald. The mixing time of the giant component of a random graph. Arxiv: math/0610459, 2006.

[5] Itai Benjamini and Elchanan Mossel. On the mixing time of a simple random walk on the super critical percolation cluster. *Probability Theory and Related Fields*, 125(3):408–420, 2003.

[6] Pietro Caputo and Alessandra Faggionato. Isoperimetric inequalities and mixing time for a random walk on a random point process. *Annals of Applied Probability*, 17(5/6):1707–1744, 2007.

[7] Pietro Caputo, Thomas M. Liggett, and Thomas Richthammer. Proof of aldous' spectral gap conjecture. Journal of the American Mathematical Society (to appear). ArXiv: 0906.1238, 2010.

[8] Colin Cooper, Alan Frieze, and Tomasz Radzik. Multiple random walks and interacting particle systems. In *Proceedings of ICALP*, pages 399–410, 2009.

[9] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible Markov chains. *Annals of Applied Probability*, 3(3):696–730, 1993.

[10] Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Annals of Applied Probability*, 6(3):695–750, 1996.

[11] Ton Dieker. Interlacings for random walks on weighted graphs and the interchange process. *SIAM Journal on Discrete Mathematics*, 24:191–206, 2010.

[12] Tzong-Yow Lee and Horng-Tzer Yau. Logarithmic Sobolev inequality for some models of random walks. *Annals of Probability*, 26(4):1855–1873, 1998.

[13] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.

[14] Thomas M. Liggett. A characterization of the invariant measures for an infinite particle system with interactions. II. *Transactions of the American Mathematical Society*, 198:201–213, 1974.

[15] Thomas M. Liggett. *Interacting Particle Systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, 1985.

[16] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*. Springer-Verlag, 1999, series=.

[17] Ravi Montenegro and Prasad Tetali. *Mathematical Aspects of Mixing Times in Markov Chains*, volume 1 of *Foundations and trends in theoretical computer science*. Now Publishers, 2006.

[18] Ben Morris. Spectral gap for the zero range process with constant rate. *Annals of Probability*, 34(5):1645–1664, 2006.

[19] Ben Morris. Improved mixing time bounds for the Thorp shuffle and L-reversal chain. *Annals of Probability*, 37(2):453–477, 2009.

[20] Benjamin Morris. The mixing time for simple exclusion. *Annals of Applied Probability*, 16(2):615–635, 2006.

[21] Benjamin Morris and Yuval Peres. Evolving sets, mixing and heat kernel bounds. *Probability Theory and Related Fields*, 133(2):245–266, 2005.

[22] Gábor Pete. A note on percolation on $\mathrm{Z}^d$: isoperimetric profile via exponential cluster repulsion. *Electronic Communications in Probability*, 13:37, 2008.

[23] Horng-Tzer Yau. Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probability Theory and Related Fields*, 109(4):507–538, 1997.