The Annals of Probability
2013, Vol. 41, No. 2, 871–913
DOI: 10.1214/11-AOP714

MIXING OF THE SYMMETRIC EXCLUSION PROCESSES IN TERMS OF THE CORRESPONDING SINGLE-PARTICLE RANDOM WALK

BY ROBERTO IMBUZEIRO OLIVEIRA[1]

IMPA

We prove an upper bound for the ε-mixing time of the symmetric exclusion process on any graph G, with any feasible number of particles. Our estimate is proportional to $\mathsf{T}_{\mathsf{RW}(G)} \ln(|V|/\varepsilon)$, 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's 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 $k \le |V|$, k-particle symmetric exclusion on G has the following informal description.

Informal description of $\mathsf{EX}(k, G)$: 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 [1], comparison arguments [8], the martingale method for log-Sobolev inequalities [12, 23] and variants of the evolving sets technology [19, 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 $(\mathbb{Z}/L\mathbb{Z})^d$ [19, 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

---

Received September 2010; revised August 2011.

[1]Supported by a *Bolsa de Produtividade em Pesquisa* and a *Pronex* project on Probability from CNPq, Brazil.

MSC2010 subject classifications. Primary 60J27, 60K35; secondary 82C22.

Key words and phrases. Symmetric exclusion, interchange process, mixing time.

denote by $\mathsf{RW}(G)$. Our result is very general, but we will see that it nearly matches previously known mixing results for $\mathsf{EX}(k, G)$ for very specific G and also gives new results in many interesting classes of examples. We will also argue that the kind of result presented here is of conceptual interest.

1.1. *The main result, and why it is interesting*. Recall that the ε-mixing-time of an irreducible continuous-time Markov chain Q on a finite set S, with transition probabilities $\{q_t(s, s')\}_{s,s'\in S,t\geq 0}$, and stationary (equilibrium) distribution π, is given by the formula

$$\mathsf{T}_Q(\varepsilon) \equiv \inf\Big\{t \geq 0 : \max_{s\in S} d_{\mathrm{TV}}(q_t(s, \cdot), \pi) \leq \varepsilon\Big\}, \tag{1.1}$$

where d_{TV} is the total-variation distance; 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 follows.

THEOREM 1.1 (Main result; proven in Section 1.4). *There exists a universal constant $C > 0$ such for all $\varepsilon \in (0, 1/2)$, all connected weighted graphs $G = (V, E, \{w_e\}_{e\in E})$ with $|V| \geq 2$ and all $k \in \{1, \dots, |V| - 1\}$,*

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

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:

$$[\textit{Heuristic assumption}] \quad \mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) \leq C_0\mathsf{T}_{\mathsf{RW}(k,G)}(\varepsilon), \qquad C_0 > 0 \text{ universal.}$$

This assumption, if at all true, is well beyond the reach of present techniques. However, it is at least plausible, given that $\mathsf{RW}(k, G)$ and $\mathsf{EX}(k, G)$ are similar.

It can be shown that $\mathsf{T}_{\mathsf{RW}(k,G)}(\varepsilon)$ and $\mathsf{T}_{\mathsf{RW}(G)}(\varepsilon/k)$ are of the same order if $\varepsilon/k \ll 1$; thus our assumption is equivalent to

$$[\textit{Heurisitic assumption}] \quad \mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) \leq C_1\mathsf{T}_{\mathsf{RW}(G)}(\varepsilon/k), \qquad C_1 > 0 \text{ universal.}$$

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 $0 < \delta < 1/2$ [1]. Applying this to our assumption, we obtain

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

Theorem 1.1 coincides with this for $k > |V|^c$, $c > 0$ a universal 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, 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 "$\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) \le C\mathsf{T}_{\mathsf{RW}(G)}(1/4)\ln(k/\varepsilon)$" is tight up to constant factors for some G (e.g., discrete tori $(\mathbb{Z}/L\mathbb{Z})^d$, d fixed [19]); therefore, in some sense Theorem 1.1 is quite close to the best that one might hope for.

Many other complex Markov chains are built from simpler processes that interact; examples appear in, for example, [1, 7, 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 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's conjecture.* Another motivation for our paper is a conjecture of Aldous's for the *interchange process*, which was recently proved in [6]. The *interchange process on G with $k \le |V|$ particles* can be informally described as follows:

Informal description of $\mathsf{IP}(k, G)$: Start with k distinct vertices of V labeled $1, 2, \dots, k$ 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 [1] implies that $\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) \le \mathsf{T}_{\mathsf{IP}(k,G)}(\varepsilon)$ for all $1 \le k \le |V| - 1$ and all $\varepsilon \in (0, 1)$.

Aldous conjectured—and Caputo et al. recently proved [6] (see also [10])—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 $\mathsf{T}_{\mathsf{IP}(k,G)}(\varepsilon)$ or $\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon)$ that can be obtained from the spectral gap are typically very loose.

Theorem 1.1 gives tighter relations between these mixing times. In the proof of the theorem, we will show that the bound claimed for $\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon)$ in the theorem statement in fact holds for $\mathsf{T}_{\mathsf{IP}(k,G)}(\varepsilon)$ whenever $k \le |V|/2$. Our proofs can be adapted to show that

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

That is, one can get a bound similar to Theorem 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 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

TABLE 1

Bounds for $\mathsf{T}_{\mathsf{EX}(k,G)}(1/4)$ *via Theorem 1.1 in examples where no previous bound was available. We take d as a fixed parameter and assume* $k \approx |V|/2$

Example	**Bound for** $\mathsf{T}_{\mathsf{EX}(k,G)}(1/4)$
$(\mathbb{Z}/L\mathbb{Z})^d$ with nearest-neighbor bonds [19]	$\|V\|^{2/d} \ln \|V\|$
Typical largest percolation cluster in $(\mathbb{Z}/L\mathbb{Z})^d$ [4, 22]	$\|V\|^{2/d} \ln \|V\|$
Typical Poisson process, in $[0, L]^d$ [5] (case $\alpha > d$)	$\|V\|^{2/d} \ln \|V\|$
Bounded-degree expanders	$\ln^2 \|V\|$
Giant component of $G_{n,c/n}$, $c > 1$ [11]	$\ln^3 \|V\|$

analyse than $\mathsf{EX}(k, G)$. The only example where we know Theorem 1.1 gives a suboptimal bound is in the case $G = (\mathbb{Z}/L\mathbb{Z})^d$ with the usual bonds, where the optimal bound, obtained by Morris [19], is of the order $L^2 \ln k$ whereas ours is about $L^2 \ln L$ (both for d fixed). Notice that this difference is only relevant for quite small k.

Table 1 presents the bounds given by Theorem 1.1 in examples where no previous bound appears explictly in the literature. The references are to papers where the mixing times of the corresponding graphs are computed. We consider only $k \approx |V|/2$ and omit constant factors.

In fairness, we note that a combination of canonical paths, log Sobolev constants and comparison arguments could in principle be applied to examples. This method is discussed in Section A in the Appendix. However, we note that:

- To the best of our knowledge, no good canonical paths bounds have been worked out for the examples in Table 1, and it might be hard or impossible to do so;
- Even if such bounds were obtained, there are natural lower bounds for how good they can be (cf. Section A), and Theorem 1.1 is at least as good as these lower bounds, up to the constants (it is actually better by a $\ln |V|$ factor in the case of expanders).

1.4. *Key steps of the proof.* Our proof of Theorem 1.1 can be broken into two main steps. We first show that $\mathsf{IP}(2, G)$ always has a mixing time comparable to $\mathsf{RW}(G)$.[2]

LEMMA 1.1. *For any weighted graph* G,

$$\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \leq 20{,}000\mathsf{T}_{\mathsf{RW}(G)}(1/4).$$

We then bootstrap the first lemma to a larger number of particles.

[2]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.

LEMMA 1.2 (Proven in Section 6.1). *There exists a universal constant $K > 0$ such that for all connected weighted graphs $G = (V, E, \{w_e\}_{e\in E})$, all $\varepsilon \in (0, 1/2)$ and all $k \in \{1, \dots, |V|/2\}$,*

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

Before we continue, we show how Theorem 1.1 easily follows from the two lemmas.

PROOF OF THEOREM 1.1. Combining Lemma 1.2 with Lemma 1.1 gives

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

where $C = 20{,}000K$. The contraction principle [1] implies

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

However, $\mathsf{EX}(k, G)$ and $\mathsf{EX}(|V| - k, G)$ are the same process with the roles of empty and occupied sites reversed. In particular, $\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) = \mathsf{T}_{\mathsf{EX}(|V|-k,G)}(\varepsilon)$ for all ε. $\square$

We now give an overview of the main ideas involved in proving the two lemmas. The proof of Lemma 1.1 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 $O(\mathsf{T}_{\mathsf{RW}(G)}(1/4))$ from *any pair* of initial states. In this case, an argument of Aldous and Fill's [1] suffices to prove Lemma 1.1 (see Proposition 4.4).

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 $\Omega(\mathsf{T}_{\mathsf{RW}(G)}(1/4))$; cf. Proposition 4.5. Intuitively, $\mathsf{IP}(2, G)$ and $\mathsf{RW}(2, G)$ are similar in the abscence of collisions, and we will use this to prove Lemma 1.1 over noneasy graphs. The negative correlation property will be crucial for this part of the argument; see Remark 4.3 for details.

The proof of Lemma 1.2 is considerably more involved. We first note that there are two methods in the literature for moving from mixing of pairs of particles to many more particles, both of which were introduced by Morris [19, 20]. The first one [20] gives bounds for walks on the symmetric group by random transpositions. Unfortunately, the method seems to require too much from the process to be useful in our general setting. Moreover, the bounds given by that method would have a factor of $\ln(|V|)\ln(1/\varepsilon)$ where ours has a $\ln(|V|/\varepsilon)$ term.

Morris's other method was introduced in his study of symmetric exclusion over $(\mathbb{Z} \setminus L\mathbb{Z})^d$ [19]. The so-called *chameleon process* features particles that change color in a way that encodes the conditional distribution of the kth particle in $\mathsf{IP}(k, G)$ given the other $k - 1$ particles. It is this method that we will successfully adapt to prove Lemma 1.2.

One way to understand Morris's construction is that it reduces the analysis of mixing to the study of pairwise collisions between particles. The analysis for $(\mathbb{Z} \setminus L\mathbb{Z})^d$ is greatly facilitated by the explicit structure of the graph, something that we lack in general. This will require certain technical modifications of Morris's construction, of which we will try to make sense with remarks in our proofs.

1.5. *Organization.* Section 2 reviews some preliminary material. Section 3 discusses $\mathsf{RW}(G)$, $\mathsf{EX}(k, G)$ and $\mathsf{IP}(k, G)$, presents their joint graphical construction and reviews the negative correlation property. Section 4 presents the proof of Lemma 1.1. The chameleon process is introduced in Section 5. It is then used to prove Lemma 1.2 in Section 6, but several lemmas are postponed to Sections 7–9. It would be pointless to describe these steps now, but Section 6.2 provides an outline of those sections. Finally, Section 10 presents some final remarks, and the Appendix contains some technical steps that are not particularly illuminating.

2. Preliminaries.

2.1. *Basic notation.* $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ is the set of nonnegative integers and $\mathbb{N}_+ \equiv \mathbb{N} \setminus \{0\}$. 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. *The elements of $(S)_k$ will always be denoted by boldface letters such as $\mathbf{x}$, with $\mathbf{x}(i)$ denoting the ith coordinate of $\mathbf{x}$.*

Notice that with these symbols,

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

A *graph* is a couple $H = (V, E)$ where $V \neq \varnothing$ 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.* $\mathcal{L}[X]$ denotes the *law* or *distribution* of the random variable X.

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

$$d_{\mathrm{TV}}(\mu, \nu) \equiv \max_{A \subset S}\big(\mu(A) - \nu(A)\big) \tag{2.2.1}$$

$$= \sup_{f:S\to[0,1]} \int f\,d\mu - \int f\,d\nu \tag{2.2.2}$$

$$= \sum_{s\in S}\big(\mu(s) - \nu(s)\big)_+ \tag{2.2.3}$$

$$= \frac{1}{2}\sum_{s\in S}|\mu(s) - \nu(s)|. \tag{2.2.4}$$

Another equivalent definition of d_{TV} is

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

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

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

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

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

We will write $\mathrm{Unif}(S)$ for the *uniform distribution* on a set $S \neq \varnothing$. 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

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

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

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

and a sequence $\{s_i\}_{i\ge 0} \subset S$ 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)],

$$\mathbb{P}(X^s_{t+h} = s' | X^s_{t'} = x_{t'}, 0 \leq t' \leq t) = \mathbb{P}(X^{x_t}_h = s') \qquad \text{almost surely.}$$

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

$$q(s, s') \equiv \lim_{\varepsilon \searrow 0} \frac{\mathbb{P}(X^s_\varepsilon = s')}{\varepsilon} \qquad [(s, s') \in S^2, s \neq s'],$$

or equivalenty by its *generator*,

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

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*, that is, chains for which for all $A \subset S$ with $A \neq \varnothing$, $S \setminus A \neq \varnothing$, there exist $a \in A$, $b \in S \setminus A$ with $q(a, b) > 0$. It is well known that such Markov chains have a unique *stationary* distribution π, that is, a distribution such that if s_* is picked according to π independently from the $(X^s_t)_{t\geq 0, s\in S}$, then $\mathcal{L}[X^{s_*}_t] = \pi$ for all $t \geq 0$. Moreover,

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

(The symbol "$\searrow$" denotes monotone convergence.) The *ε-mixing time of Q* is thus defined as in the Introduction,

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

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

PROPOSITION 2.1 ([13], equation (4.36), page 55). *Let Q be a Markov chain on finite state space S. Then for all* $0 < \varepsilon < 1/2$,[3]

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

PROPOSITION 2.2 ([1], Lemma 7 in Chapter 4). *Let Q be a Markov chain on finite state space S with symmetric transition rates. Then π is uniform over S and moreover, for all* $0 < \varepsilon < 1/2$ *and* $t \geq 2\mathsf{T}_Q(\varepsilon)$,

$$\mathbb{P}(X^s_t = s') \geq \frac{(1 - 2\varepsilon)^2}{|S|},$$

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

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

We also make the following convenient notational convention.

NOTATIONAL CONVENTION 2.2. *By definition, for any càdlag path* $(x_t)_{t\geq 0}$ *there exists a divergent sequence* $t_0 = 0 < t_1 < t_2 < \cdots$ *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*,

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

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 [1] 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. It will be useful to 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 also write $f_e(A) = \{f_e(a) : a \in A\}$ and $f_e(\mathbf{x}) = (f_e(\mathbf{x}(i)))_{i=1}^k$ for $A \in \binom{V}{k}$ and $\mathbf{x} \in (V)_k$ (resp.).

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

$$q(u, v) \equiv \begin{cases} w_e, & \text{if } f_e(u) = v; \\ 0, & \text{otherwise} \end{cases} \qquad [(u, v) \in (V)_2].$$

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 these process are also symmetric, it follows that the stationary distribution of $\mathsf{RW}(k, G)$ is $\mathrm{Unif}(V^k)$ for all $k \in \mathbb{N}_+$.

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

$$q^{\{k\}}(A, B) \equiv \begin{cases} w_e, & \text{if } f_e(A) = B; \\ 0, & \text{otherwise} \end{cases} \qquad \left[(A, B) \in \left(\binom{V}{k}\right)_2\right].$$

The transition rates are again symmetric, and the stationary distribution of $\mathsf{EX}(k, G)$ is $\mathrm{Unif}(\binom{V}{k})$.

The k-particle interchange process on G, denoted by $\mathsf{IP}(k, G)$, has state space $(V)_k$. The transition rates of $\mathsf{IP}(k, G)$ are given by

$$q^{(k)}(\mathbf{x}, \mathbf{y}) \equiv \begin{cases} w_e, & \text{if } f_e(\mathbf{x}) = \mathbf{y}; \\ 0, & \text{otherwise} \end{cases} \qquad [(\mathbf{x}, \mathbf{y}) \in ((V)_k)_2].$$

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

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. For brevity, we omit all proofs in this subsection.

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

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

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

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 $\mathcal{P} \cap (t, s] = \varnothing$, $I_{(t,s]}$ is the identity map on V. If, on the other hand,

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

we set $I_{(t,s]} = f_{e_n} \circ f_{e_{n-1}} \circ \cdots \circ f_{e_m}$; that is, $I_{(t,s]}$ 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 3.1. 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 3.1 (Proof omitted). *For all* $0 \leq t \leq s \leq r$, $I_{(t,r]} = I_{(s,r]} \circ I_{(t,s]}$.

PROPOSITION 3.2 (Proof omitted). *For all* $0 \leq t \leq s < +\infty$, $\mathcal{L}[I_{(t,s]}] = \mathcal{L}[I^{-1}_{(t,s]}]$.

PROPOSITION 3.3 (Proof omitted). *Let* $0 \le t_0 < t_1 < t_2 < \cdots < t_k$. *Then the maps* $I_{(t_{i-1},t_i]}$, $1 \le i \le k$, *are independent.*

NOTATIONAL CONVENTION 3.1. *We "lift" the random maps* $I_{(t,s]}$ *to permutations of* $\binom{V}{k}$ *and* $(V)_k$, *which we also denote by* $I_{(t,s]}$:

$$I_{(t,s]}(A) \equiv \{I_{(t,s]}(a) : a \in A\} \qquad \left[A \in \binom{V}{k}\right],$$
$$I_{(t,s]}(\mathbf{x}) \equiv \bigl(I_{(t,s]}(\mathbf{x}(1)), I_{(t,s]}(\mathbf{x}(2)), \ldots, I_{(t,s]}(\mathbf{x}(k))\bigr) \qquad [\mathbf{x} \in (V)_k].$$

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

The key property of the graphical construction follows:

PROPOSITION 3.4 (Proof omitted). *Let* $t_0 \ge 0$. *Then*:

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

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 [3, 14, 15].

LEMMA 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* $\mathbf{u} \in (V)_2$—*that is, for all distinct* $\mathbf{u}(1), \mathbf{u}(2) \in V$—*we have*

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

Using the construction in the previous section, we can write the above inequality as

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

The following is then immediate from Proposition 3.2.

COROLLARY 3.1 (Proof omitted). *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* $A \subset V$,

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

4. The dynamics of pairs of particles. The goal of this section is to prove Lemma 1.1. 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. *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. The first one is obvious, for example, from the graphical construction.

PROPOSITION 4.1 (Proof omitted). *For* $i = 1, 2$, $\mathcal{L}[\mathbf{x}_t^R(i)] = \mathcal{L}[\mathbf{x}_t^I(i)]$.

The next proposition is a direct consequence of (2.2.5).

PROPOSITION 4.2 (Proof omitted). *The mixing times of* $\mathsf{RW}(2, G)$ *satisfy* $\mathsf{T}_{\mathsf{RW}(2,G)}(\varepsilon) \leq \mathsf{T}_{\mathsf{RW}(G)}(\varepsilon/2)$.

PROPOSITION 4.3. *Let* $k \in \mathbb{N}$ *be given. Then* $\mathsf{T}_{\mathsf{RW}(2,G)}(2^{-k}) \leq (k + 1) \times \mathsf{T}_{\mathsf{RW}(G)}(1/4)$.

PROOF. Follows from the previous proposition combined with Proposition 2.1. $\square$

The next lemma has the following meaning. Suppose t is so large that $\mathbf{x}_t^R$ is close to equilibrium. In this case, $\mathbb{E}[\phi(\mathbf{x}_t^R)]$ is close to the uniform average of ϕ over V^2, for all mappings $0 \leq \phi \leq 1$. The lemma shows that $\mathbb{E}[\phi(\mathbf{x}_t^I)]$ cannot be much larger than that average. This will require the negative correlation property; cf. Corollary 3.1.

LEMMA 4.1. *Let* $\phi : V^2 \to [0, 1]$. *Then*

$$\forall \varepsilon \in (0, 1/16), \forall t \geq \mathsf{T}_{\mathsf{RW}(G)}(\varepsilon), \forall \mathbf{x} \in (V)_2 \qquad \mathbb{E}[\phi(\mathbf{x}_t^I)] \leq 8\sqrt{\varepsilon} + 9 \sum_{\mathbf{v} \in V^2} \frac{\phi(\mathbf{v})}{|V|^2}.$$

PROOF. Define the "good set" of all $a \in V$ with nearly uniform probability

$$\text{Good} \equiv \left\{a \in V : \max_{i=1,2}\left|\mathbb{P}(\mathbf{x}_t^R(i) = a) - \frac{1}{|V|}\right| \le \frac{2\sqrt{\varepsilon}}{|V|}\right\}.$$

We will show toward the end of the proof that

$$\mathbb{P}(\mathbf{x}_t^I \notin \text{Good}^2) \le 8\sqrt{\varepsilon}, \tag{4.1.1}$$

which (since $0 \le \phi \le 1$) implies

$$\mathbb{E}[\phi(\mathbf{x}_t^I)\mathbb{I}_{(V)_2\setminus\text{Good}^2}(\mathbf{x}_t^I)] \le 8\sqrt{\varepsilon}. \tag{4.1.2}$$

On the other hand, notice that

$$\begin{aligned}
\mathbb{E}[\phi(\mathbf{x}_t^I)\mathbb{I}_{\text{Good}^2}(\mathbf{x}_t^I)] &= \sum_{\mathbf{a}\in(\text{Good}^2)\cap(V)_2} \mathbb{P}(\mathbf{x}_t^I = (\mathbf{a}(1), \mathbf{a}(2)))\phi(\mathbf{a}) \\
&\le \sum_{\mathbf{a}\in(\text{Good}^2)\cap(V)_2} \mathbb{P}\left(\bigcap_{i=1}^2 \{\mathbf{x}_t^I(i) \in \{\mathbf{a}(1), \mathbf{a}(2)\}\}\right)\phi(\mathbf{a}) \\
(\text{Cor. } 3.1) &\le \sum_{\mathbf{a}\in(\text{Good}^2)\cap(V)_2} \prod_{i=1}^2 \mathbb{P}(\{\mathbf{x}_t^I(i) \in \{\mathbf{a}(1), \mathbf{a}(2)\}\})\phi(\mathbf{a}) \\
(\text{Prop. } 4.1) &= \sum_{\mathbf{a}\in(\text{Good}^2)\cap(V)_2} \prod_{i=1}^2 \mathbb{P}(\{\mathbf{x}_t^R(i) \in \{\mathbf{a}(1), \mathbf{a}(2)\}\})\phi(\mathbf{a}) \\
[\mathbf{a}(i) \in \text{Good}] &\le \sum_{\mathbf{a}\in(\text{Good}^2)\cap(V)_2} \left(\frac{2 + 4\sqrt{\varepsilon}}{|V|}\right)^2 \phi(\mathbf{a}) \\
(\sqrt{\varepsilon} \le 1/4) &\le 9\sum_{\mathbf{a}\in V^2} \frac{\phi(\mathbf{a})}{|V|^2}.
\end{aligned}$$

Combining this with (4.1.2) finishes the proof, except for (4.1.1). To prove that, we let $\text{Bad} = V \setminus \text{Good}$. Notice that

$$\frac{\sqrt{\varepsilon}|\text{Bad}|}{|V|} \le \sum_{a\in V} \frac{1}{2}\left\{\left|\mathbb{P}(\mathbf{x}_t^R(1) = a) - \frac{1}{|V|}\right| + \left|\mathbb{P}(\mathbf{x}_t^R(2) = a) - \frac{1}{|V|}\right|\right\}$$

as each $a \in \text{Bad}$ contributes at least $\sqrt{\varepsilon}/|V|$ to the sum. But the RHS equals

$$d_{\text{TV}}(\mathcal{L}[\mathbf{x}_t^R(1)], \text{Unif}(V)) + d_{\text{TV}}(\mathcal{L}[\mathbf{x}_t^R(2)], \text{Unif}(V)) \le 2\varepsilon$$

since $t \ge \mathsf{T}_{\mathsf{RW}(G)}(\varepsilon)$. We deduce

$$\frac{\sqrt{\varepsilon}|\text{Bad}|}{|V|} \le 2\varepsilon \quad \text{or equivalently} \quad |\text{Good}| \ge (1 - 2\sqrt{\varepsilon})|V|.$$

Moreover, $\mathbb{P}(\mathbf{x}_t^R(i)=a) \geq (1-2\sqrt{\varepsilon})|V|^{-1}$ for all $a \in \text{Good}$, hence

$$\mathbb{P}(\mathbf{x}_t^R(i) \in \text{Good}) \geq \frac{|\text{Good}|}{|V|}(1-2\sqrt{\varepsilon}) \geq (1-2\sqrt{\varepsilon})^2 \geq 1-4\sqrt{\varepsilon}.$$

Inequality (4.1.1) now follows from

$$\begin{aligned} \mathbb{P}(\mathbf{x}_t^I \notin \text{Good}^2) &\leq \mathbb{P}(\mathbf{x}_t^I(1) \notin \text{Good}) + \mathbb{P}(\mathbf{x}_t^I(2) \notin \text{Good}) \\ (\text{Proposition 4.1}) &= \mathbb{P}(\mathbf{x}_t^R(1) \notin \text{Good}) + \mathbb{P}(\mathbf{x}_t^R(2) \notin \text{Good}) \leq 8\sqrt{\varepsilon}. \qquad \square \end{aligned}$$

4.2. *When collisions are nearly as fast as mixing.* Recalling Notational convention 4.1, 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 $\mathbf{x}_{t_0}^R(1) = \mathbf{x}_{t_0}^R(2)$ (this is a.s. finite by ergodicity). We will also write

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

for the time until the first meeting after t (this is a "time-shifted" meeting time).

The following definition will be crucial for our analysis.

DEFINITION 4.1. We say that a weighted graph G is easy if

$$\sup_{\mathbf{x} \in V^2} \mathbb{P}(M(\mathbf{x}) > 20{,}000\mathsf{T}_{\mathsf{RW}(G)}(1/4)) \leq 1/8.$$

We note that all long enough paths and cycles are examples of easy graphs. Noneasy graphs include $(\mathbb{Z}/L\mathbb{Z})^d$ for $d \geq 2$ fixed and L sufficiently large, as well as large expander graphs. The next proposition proves Lemma 1.1 for all easy graphs via a coupling argument due to Aldous and Fill.

PROPOSITION 4.4. *Lemma 1.1 holds for all easy weighted graphs.*

PROOF SKETCH. Given G, Aldous and Fill [1, Chapter 14, Section 5] construct a coupling of $\mathsf{IP}(|V|,G)$ started from two different states $\mathbf{u}, \mathbf{v}$. Letting $\{\mathbf{u}_t^I, \mathbf{v}_t^I\}_{t\geq 0}$ denote the coupled trajectories, the following property holds: for each $1 \leq i \leq |V|$, $\mathbf{u}_t^I(i), \mathbf{v}_t^I(i)$ behave as independent random walks up to their first meeting time, which we denote by M_i. After this time M_i, $\mathbf{u}_t^I(i) = \mathbf{v}_t^I(i)$, that is, the two processes move together. This implies

$$\begin{aligned} \forall t \geq 0 \qquad d_{\mathrm{TV}}(\mathcal{L}[\mathbf{u}_t^I], \mathcal{L}[\mathbf{v}_t^I]) \leq \mathbb{P}(\mathbf{u}_t^I \neq \mathbf{v}_t^I) &\leq \sum_{i=1}^{|V|} \mathbb{P}(\mathbf{u}_t^I(i) \neq \mathbf{v}_t^I(i)) \\ &\leq \sum_{i=1}^{|V|} \mathbb{P}(M_i > t). \end{aligned}$$

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 $\{\mathbf{x}_t^I, \mathbf{y}_t^I\}_{t\geq 0}$ denotes the coupled trajectories, we have

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

Now both M_1 and M_2 are the meeting times of independent random walkers on G, which shows that

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

For $t = 20{,}000\mathsf{T}_{\mathsf{RW}(G)}(1/4)$ and G easy, the RHS is $\leq 1/4$. By convexity, this implies that

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

In other words, $\mathsf{T}_{\mathsf{IP}(2,G)}(1/4) \leq 20{,}000\mathsf{T}_{\mathsf{RW}(G)}(1/4)$. $\square$

REMARK 4.1. Aldous and Fill's argument actually proves Theorem 1.1 for all easy graphs; see [1], Chapter 14, Section 5 for details.

4.3. *Long time to meet in noneasy 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 4.5. *Assume $G = (V, E, \{w_e\}_{e\in E})$ is not easy. Then*

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

REMARK 4.2. In general we cannot guarantee that $\mathbb{P}(M(\mathbf{v}) < 20 \times \mathsf{T}_{\mathsf{RW}(G)}(1/4))$ is uniformly small over all $\mathbf{v} \in (V)_2$. In particular, the probability of collision from adjacent $\mathbf{v}(1), \mathbf{v}(2)$ might be much greater than the above bound.

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

$$\mathbb{P}\big(M(\mathbf{x}) > 20{,}000T\big) > 1/8. \tag{4.3.1}$$

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

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

The conditional probability in the RHS equals $\mathbb{P}(M(\mathbf{y}) > 40T)$ for $\mathbf{y} = \mathbf{x}^R_{40(k-1)T}$, hence

$$\mathbb{P}\big(M(\mathbf{x}) > 40kT\big) \le \Big(\sup_{\mathbf{y}\in V^2} \mathbb{P}\big(M(\mathbf{y}) > 40T\big)\Big)\mathbb{P}\big(M(\mathbf{x}) > 40(k-1)T\big)$$
$$(\ldots\text{induction}\ldots) \le \Big(\sup_{\mathbf{y}\in V^2} \mathbb{P}\big(M(\mathbf{y}) > 40T\big)\Big)^k.$$

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

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

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

$$\frac{497}{500} \le \mathbb{P}\big(M(\mathbf{y}) > 40T\big) \le \mathbb{P}\big(M_{\ge 20T}(\mathbf{y}) > 20T\big)$$
$$= \sum_{\mathbf{v}\in V^2} \mathbb{P}(\mathbf{y}^R_{20T} = \mathbf{v})\mathbb{P}\big(M(\mathbf{v}) > 20T\big).$$

Moreover, by (2.2.2),

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

Hence

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

We finish by noting that, by Proposition 4.3, $20T \ge \mathsf{T}_{\mathsf{RW}(2,G)}(2^{-19})$, hence

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

and therefore

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

□

4.4. *If meeting takes a long time*, $\mathsf{IP}(2, G)$ *and* $\mathsf{RW}(2, G)$ *are similar.* We have just shown that the meeting is unlikely to be smaller than $20\mathsf{T}_{\mathsf{RW}(G)}(1/4)$ from most initial states. We now show that $\mathsf{IP}(2, G)$ is similar to $\mathsf{RW}(2, G)$ until the first meeting time.

PROPOSITION 4.6. *For any* $\mathbf{x} \in (V)_2$ *and* $s \geq 0$,

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

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

COROLLARY 4.1. *For any* $\mathbf{x}, \mathbf{y} \in (V)_2$ *and* $s \geq 0$,

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

PROOF OF PROPOSITION 4.6. 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 Δ^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 $e = \{\mathbf{x}(1), \mathbf{x}(2)\}$,

$$\begin{cases} ((\mathbf{x}, \mathbf{x}), (f_e(\mathbf{x}), (\mathbf{x}(1), \mathbf{x}(1)))) = w_e, \\ q((\mathbf{x}, \mathbf{x}), (\mathbf{x}, (\mathbf{x}(2), \mathbf{x}(2)))) = w_e. \end{cases}$$

  - *Transition rule* 2.3: All other potential transitions have rate 0.

Inspection of the marginals reveals that this indeed gives a coupling of $\{\mathbf{x}_t^R\}_{t\geq 0}$ and $\{\mathbf{x}_t^I\}_{t\geq 0}$ when started from an initial state $(\mathbf{x}, \mathbf{x}) \in \Delta$. Moreover, the two processes can only differ after a transition has occurred according to rule 2.2. The first time when this happens is precisely the first meeting time of $\{\mathbf{x}_t^R\}_{t\geq 0}$. $\square$

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 1.1.

PROOF OF LEMMA 1.1. The case of easy graphs is covered by Proposition 4.4, 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 $\{\mathbf{x}^I_t\}_{t\geq 0}$, $\{\mathbf{y}^I_t\}_{t\geq 0}$ are defined over the same probability space,

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

Maximizing over A yields

$$\begin{aligned}&d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}^I_{40T}],\mathcal{L}[\mathbf{y}^I_{40T}])\\&\quad\leq \mathbb{E}\big[d_{\mathrm{TV}}\big(\mathbb{P}(\mathbf{x}^I_{40T}\in\cdot|\mathbf{x}^I_{20T}),\mathbb{P}(\mathbf{y}^I_{40T}\in\cdot|\mathbf{y}^I_{20T})\big)\big].\end{aligned}\tag{4.5.1}$$

By the Markov property and Corollary 4.1,

$$\begin{aligned}&d_{\mathrm{TV}}\big(\mathbb{P}(\mathbf{x}^I_{40T}\in\cdot|\mathbf{x}^I_{20T}=\mathbf{v}),\mathbb{P}(\mathbf{y}^I_{40T}\in\cdot|\mathbf{y}^I_{20T}=\mathbf{w})\big)\\&\quad=d_{\mathrm{TV}}(\mathcal{L}[\mathbf{v}^I_{20T}],\mathcal{L}[\mathbf{w}^I_{20T}])\\&\quad\leq \mathbb{P}\big(M(\mathbf{v})\leq 20T\big)+\mathbb{P}\big(M(\mathbf{w})\leq 20T\big)+d_{\mathrm{TV}}(\mathcal{L}[\mathbf{v}^R_{20T}],\mathcal{L}[\mathbf{w}^R_{20T}]).\end{aligned}$$

Proposition 4.3 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

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

where $\phi(\mathbf{z})=\mathbb{P}(M(\mathbf{z})\leq 20T)$. Notice that $0\leq\phi\leq 1$. We may apply Lemma 4.1 and the fact that $20T\geq \mathsf{T}_{\mathsf{RW}(G)}(2^{-20})$ (cf. Proposition 2.1) to deduce

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

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

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

Finally, we use the fact that G is not easy, combined with Proposition 4.5, 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 1.1 for noneasy graphs. $\square$

REMARK 4.3. The first inequality in (4.5.3) follows from Lemma 4.1, which is a consequence of the negative correlation property; cf. Lemma 3.1 and Corollary 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's paper [19]. 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 $\mathcal{P} = \{\tau_1 \le \tau_2 \le \tau_3 \le \cdots\} \subset [0, +\infty)$ with rate $2W$.
(2) An i.i.d. sequence of E-valued random variables $\{e_n\}_{n\in\mathbb{N}}$, with $\mathbb{P}(e_n = e) = w_e/W$.
(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, and 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 $\mathcal{P} \cap (t, s] = \varnothing$, $I_{(t,s]}$ is the identity map, as before. Otherwise,

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

and we set

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

The *thinning property* of the Poisson process implies that $\{\tau_n : c_n = 1\}$ is a Poisson process with rate W. One can use this to show that:

PROPOSITION 5.1 (Proof omitted). *The joint distribution of the maps $I_{(t,s]}$, $0 \le t < s < +\infty$, is the same as in Section 3.2.*

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 $\mathbf{O}(\mathbf{y}) \equiv \{\mathbf{y}(1), \dots, \mathbf{y}(k-1)\}$ 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

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

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* (resp.). Notice that any vertex $v \in V$ will belong to one of these color classes.

The evolution of the process from initial state $(\mathbf{z}, R, P, W)$ will be denoted by

$$\{(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C)\}_{t \geq 0}.$$

By definition, this process will only be updated at the times τ_n ($n \in \mathbb{N}$) given by the Poisson process and at deterministic times $2iT$, $i \in \mathbb{N}$. Moreover, the updates at times τ_n are of different kinds depending on whether $\tau_n \in ((2i-2)T, (2i-1)T]$ for some $i \in \mathbb{N}_+$, or $\tau_n \in ((2i-1)T, 2iT]$ for some $i \in \mathbb{N}_+$. Finally, we will *define* for convenience,

$$(\mathbf{z}_{0_-}^C, R_{0_-}^C, P_{0_-}^C, W_{0_-}^C) = (\mathbf{z}, R, P, W)$$

and will allow an instantaneous change at time $t = 0$: that is,

$$\text{it might happen that } (\mathbf{z}_0^C, R_0^C, P_0^C, W_0^C) \neq (\mathbf{z}, R, P, W).$$

The three update rules are described in Box 5.1.

REMARK 5.1. Technically, this process is *not* càdlàg, as it changes at time 0. We will nevertheless continue to use t_- (cf. Notational convention 2.2) with the proviso for $t = 0$ that we have just described.

REMARK 5.2. We briefly note that our chamaleon process is more complicated than Morris's process [19]. In brief: his process does not have constant-color phases and will depink right when the number of pink particles exceeds the minimum of red and white. The second difference is a matter of convenience, but the first one will be fundamental at key steps of our argument.

5.3. *Two basic properties*. The next two results will be useful later on. We only sketch the proofs.

LEMMA 5.1. *Let*

$$(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i) = \textit{the value of } (\mathbf{z}_{2iT_-}^C, R_{2iT_-}^C, P_{2iT_-}^C, W_{2iT_-}^C) \qquad (i \in \mathbb{N}).$$

Then $\{(\hat{\mathbf{z}}_i, \hat{R}_i, \hat{P}_i, \hat{W}_i)\}_{i \in \mathbb{N}}$ *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 5.1 The three kinds of updates in the chameleon process.

- Constant-color phases: If $t = \tau_n \in ((2i-2)T, (2i-1)T]$ for some $i \in \mathbb{N}_+$, update

$$(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C) = (f_{e_n}^{c_n}(\mathbf{z}_{t-}^C), f_{e_n}^{c_n}(R_{t-}^C), f_{e_n}^{c_n}(P_{t-}^C), f_{e_n}^{c_n}(W_{t-}^C)). \tag{5.2.2}$$

  That is, the states of the endpoints of e_n are flipped if $c_n = 1$, and nothing happens if $c_n = 0$.
- Color-changing phases. If $t = \tau_n \in ((2i-1)T, 2iT]$, for $i \in \mathbb{N}_+$, update as above *unless*:

  1. $e_n = \{w, r\}$ has a white endpoint $w \in W_{t-}^C$ and a red endpoint $r \in R_{t-}^C$;
  2. $|P_{t-}^C| < \min\{|R_{t-}^C|, |W_{t-}^C|\}$.

  If (1) and (2) hold, r and w both become pink, and we call t a pinkening time.

$$(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C) = (\mathbf{z}_{t-}^C, R_{t-}^C \setminus \{r\}, P_{t-}^C \cup \{r, w\}, W_{t-}^C \setminus \{w\}). \tag{5.2.3}$$

- Depinking times. If $t = 2iT$ with $i \in \mathbb{N}$ ($t = 0$ or t lies at the end of a color-changing phase) *and* $|P_{t-}^C| \geq \min\{|W_{t-}^C|, |R_{t-}^C|\}$ (more pink than either white or red), flip a fair coin d_i, and make all pink particles become red or white depending on whether d_i comes out heads or tails (resp.).

$$(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C) = \begin{cases} (\mathbf{z}_{t-}^C, R_{t-}^C \cup P_{t-}^C, \varnothing, W_{t-}^C), & d_i = 1; \\ (\mathbf{z}_{t-}^C, R_{t-}^C, \varnothing, W_{t-}^C \cup P_{t-}^C), & d_i = 0. \end{cases} \tag{5.2.4}$$

PROOF SKETCH. Markovianity and time-homogeneity are obvious. To prove the stopping time property, it suffices to check that (setting $D_0 = 0$),

$$\forall j > 0 \qquad \frac{D_j}{2T} = \inf\left\{i > \frac{D_{j-1}}{2T} : |\hat{P}_i| \geq \min\{|\hat{R}_i|, |\hat{W}_i|\}\right\},$$

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

LEMMA 5.2. *Suppose $(\mathbf{z}_{2iT}^C, R_{2iT}^C, P_{2iT}^C, W_{2iT}^C)$ is the state of the chameleon process at time $2iT$ (i.e., at the beginning of a constant-color phase). Then*

$$(\mathbf{z}_{(2i+1)T}^C, R_{(2i+1)T}^C, P_{(2i+1)T}^C, W_{(2i+1)T}^C) = (I(\mathbf{z}_{2iT}^C), I(R_{2iT}^C), I(P_{2iT}^C), I(W_{2iT}^C)),$$

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

PROOF. By inspection. $\square$

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

NOTATIONAL CONVENTION 5.1. *$\mathbf{x} = (\mathbf{x}(1), \dots, \mathbf{x}(k)) \in (V)_k$ is represented as a pair $(\mathbf{z}, x)$, where $\mathbf{z} = (\mathbf{x}(1), \dots, \mathbf{x}(k-1)) \in (V)_{k-1}$ and $x = \mathbf{x}(k) \in V \setminus \mathbf{O}(\mathbf{z})$. [Notice that $\mathbf{x}_t^I = (\mathbf{z}_t^I, x_t^I)$ for all $t \geq 0$.]*

PROPOSITION 5.2 (Proof omitted). *Given an initial state $\mathbf{x} = (\mathbf{z}, x) \in (V)_k$ for $\mathsf{IP}(k, G)$, set $R = \{x\}$, $P = \varnothing$ and $W = V \setminus (\mathbf{O}(\mathbf{z}) \cup \{x\})$. Consider the interchange process $\{\mathbf{x}_t^I = (\mathbf{z}_t^I, x_t^I)\}_{t \geq 0}$ started from state $\mathbf{x}$ and the chameleon process $\{(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C)\}_{t \geq 0}$ started from configuration $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$. Then*

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

where

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

This is almost identical (up to changes in notation) to [19, Lemma 1], and we omit its proof. It will be 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. From 2 to *k* particles via the chameleon process. In this section we present the proof of Lemma 1.2, modulo several lemmas about the chameleon process that we will prove later. We then outline the remainder of the paper.

6.1. *Proof of Lemma 1.2.*

PROOF. 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

$$(6.1.1) \qquad \mathsf{ink}_t \equiv \sum_{v \in V} \mathsf{ink}_t(v) = |R_t^C| + \frac{|P_t^C|}{2} \qquad (t \geq 0).$$

We note for later reference that

$$(6.1.2) \qquad \mathsf{ink}_t \equiv \sum_{v \in V \setminus \mathbf{O}(\mathbf{z}_t^I)} \mathsf{ink}_t(v)$$

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

We have argued in Proposition 5.2 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 $(\mathbf{z}, \{x\}, \varnothing, V \setminus (\mathbf{O}(\mathbf{z}) \cup \{x\}))$. Letting $\mathsf{ink}_t^{\mathbf{x}}$ denote the value of ink_t in that chameleon process, we will show that:

LEMMA 6.1 (Proven in Section 8.2). *The following inequality holds for all* $1 \le k \le |V|-1$:

$$\sup_{\mathbf{x}\in(V)_k} d_{\mathrm{TV}}\big(\mathcal{L}[\mathbf{x}_t^I], \mathrm{Unif}((V)_k)\big) \le 2k \sup_{\mathbf{x}\in(V)_k} \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{x}}}{|V|-k+1}\middle|\mathsf{Fill}\right]$$

where

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

The main goal is to bound the expected value in the RHS of the inequality in Lemma 6.1. Fix some $\mathbf{x} \in (V)_k$, and let $D_j(\mathbf{x})$ denote the jth depinking time for the chameleon process corresponding to $\mathbf{x}$. 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 that there are infinitely many depinking times, that is, there are infinitely many times of the form $2iT$ at which the number of pink particles is at least as large as the minimum of the numbers of white and red. The definition of the chameleon process implies that $\mathsf{ink}_t^{\mathbf{x}}$ can only change at depinking times, hence for any $t \ge 0$ $\mathsf{ink}_t^{\mathbf{x}} = 1$ if $t < D_1(\mathbf{x})$ and $\mathsf{ink}_t^{\mathbf{x}} = \widehat{\mathsf{ink}}_j^{\mathbf{x}}$ if $D_j(\mathbf{x}) \le t < D_{j+1}(\mathbf{x})$ for some j. We deduce that

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

Taking expectations, we see that the RHS of the inequality in Lemma 6.1 is at most

$$2k \sup_{\mathbf{x}\in(V)_k}\left\{\sum_{m\ge j}\mathbb{E}\left[1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V|-k+1}\middle|\mathsf{Fill}\right] + \mathbb{P}\big(D_j(\mathbf{x}) \ge t|\mathsf{Fill}\big)\right\}. \tag{6.1.3}$$

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

PROPOSITION 6.1 (Proven in Section B). *For all* $\ell \ge 1$ *and* $\mathbf{x} \in (V)_k$,

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

We thus have

$$\begin{aligned} &2k \sup_{\mathbf{x}\in(V)_k}\left\{\sum_{m\ge j}\mathbb{E}\left[1 - \frac{\widehat{\mathsf{ink}}_m^{\mathbf{x}}}{|V|-k+1}\middle|\mathsf{Fill}\right] + \mathbb{P}\big(D_j(\mathbf{x}) \ge t|\mathsf{Fill}\big)\right\} \\ &\quad \le C_2|V|^{3/2}e^{-c_1 j} + 10k \sup_{\mathbf{x}\in(V)_k} \mathbb{P}\big(D_j(\mathbf{x}) \ge t|\mathsf{Fill}\big), \end{aligned} \tag{6.1.4}$$

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

Bounding $\mathbb{P}(D_j(\mathbf{x}) \geq t|\mathsf{Fill})$ 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 (Proven in Section 9.3). *There exist universal constants* $C_3, C_4 > 0$, *such that if* $|V| \geq 300$, $T \geq C_3 \mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$ *and* $k/|V| \leq 1/2$, *then*

$$\forall \mathbf{x} \in (V)_k,\ \forall j \in \mathbb{N}: \qquad \mathbb{E}\big[e^{D_j(\mathbf{x})/(C_4 T)}|\mathsf{Fill}\big] \leq e^j.$$

If $|V| \geq 300$ Markov's inequality allows one to deduce that, for yet another universal constant $L \equiv C_3 C_4$,

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

Plugging this into (6.1.4) and Lemma 6.1, we obtain

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

Since this inequality holds for all j, we can take

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

and obtain

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

with $K_0 > 0$ universal. Comparing with the definition of mixing time in (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 $|V| < 300$—that is, $|V|$ bounded by a universal constant—can be dealt with in several ways. For example, one may use the result of Caputo et al. [6] 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 $\mathsf{T}_{\mathsf{IP}(k,G)}(\varepsilon)$ in terms of its spectral gap (see, e.g., [17] for these standard bounds). Alternatively, one may use Aldous and Fill's analysis (see Remark 4.1) together with the inequality

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

which one can prove via Proposition 2.2 and a few simple calculations. $\square$

6.2. *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 ink. The proof of Proposition 6.1 on the decay of $\mathbb{E}[1-\widehat{\mathsf{ink}}_\ell^{\mathbf{x}}/(|V|-k+1)|\mathsf{Fill}]$, presented in the Appendix (see Section B), relies on results from this section.
(2) Section 8 contains the proof of Lemma 6.1, which is based on an auxiliary result on conditional distributions (Lemma 8.1).
(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 jth depinking time. This leads to the key Lemma 6.2, proven in Section 9.3.

7. A miscellany of facts on ink. In this section we prove several facts we will need about the quantity ink_t introduced in (6.1.1). We will use the same notation introduced in the proof of Lemma 1.2 (cf. Section 6.1):

(1) $\mathbf{x}\in(V)_k$ is some fixed state;
(2) $(\mathbf{z},R,P,W)=(\mathbf{z},\{x\},\varnothing,V\setminus(\mathbf{O}(\mathbf{z})\cup\{x\}))\in\mathcal{C}_k(V)$ is the initial state corresponding to $\mathbf{x}$ in the sense of Proposition 5.2;
(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 jth depinking time for this process;
(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. Most proofs in this section follow by inspection, so we will be quite brief.

In principle the total number of number of depinking times could be finite. We begin by showing that this is not the case.

PROPOSITION 7.1. *The number of depinking times is almost surely infinite.*

REMARK 7.1. Notice that this only works because our *definition* of a depinking time allows for "trivial depinking times" where there are only red and black (or only white and black) particles left. This was noted in Box 5.1.

PROOF OF PROPOSITION 7.1. We use the following simple fact (proof omitted): there exists some $\delta>0$ such that each color-changing phase that starts with $\min\{|R_t^C|,|W_t^C|\}>0$ will have a pinkening with probability $\geq\delta$, *regardless of the past*. This implies that, given $s\geq 0$, the values of $|P_t^C|$ for $t\in[s,+\infty)$ will have a strictly positive probability of increasing by the end of each color-changing phase, at least until $|P_t^C|\geq\min\{|R_t^C|,|W_t^C|\}$. Since $|P_t^C|$ can only decrease at depinking steps, this shows that $|P_t^C|$ must continue to increase until $|P_t^C|\geq\min\{|R_t^C|,|W_t^C|\}$, and the next time of the $2iT$ will be a depinking time. □

The next result follows by inspection.

PROPOSITION 7.2 (Proof omitted). $0 \le \widehat{\mathsf{ink}}_j \le |V| - k + 1$ *for all* $j \in \mathbb{N}$, *a.s.*

We now compute the amount of change of ink in each step.

PROPOSITION 7.3. *For* $j \in \mathbb{N}$, $\widehat{\mathsf{ink}}_{j+1} \in \{\widehat{\mathsf{ink}}_j + \Delta(\widehat{\mathsf{ink}}_j), \widehat{\mathsf{ink}}_j - \Delta(\widehat{\mathsf{ink}}_j)\}$ *a.s., where*

$$\Delta(r) \equiv \left\lceil \frac{\min\{r, |V| - k + 1 - r\}}{3} \right\rceil \qquad (r \in \mathbb{N}). \tag{7.0.1}$$

Moreover, conditionally on $\{\widehat{\mathsf{ink}}_\ell\}_{\ell=0}^{j}$, *each possibility is equally likely.*

PROOF. Box 5.1 shows that there are no pink particles left in the system after depinking is performed. This implies that $\widehat{\mathsf{ink}}_j = \mathsf{ink}_{D_j} = |R_{D_j}^C|$. Moreover, since the total number of nonblack particles is $|V| - k + 1$, there must be $\widehat{\mathsf{ink}}_j$ red and $|V| - k + 1 - \widehat{\mathsf{ink}}_j$ white particles at time D_j.

A pinkening step decreases the number of red and white particles by 1 each and increases the number of pink particles by 2. However, no pinkenings are performed if the number of pink particles is at least the number of red or the number of white particles. In other words, the number of pinkening steps until the next depinking is precisely the *smallest* p satisfying $2p \ge \widehat{\mathsf{ink}}_j - p$ or $2p \ge |V| - k - 1 - \widehat{\mathsf{ink}}_j - p$, which is $p = \Delta(\widehat{\mathsf{ink}}_j)$ for Δ defined in (7.0.1).

At the depinking step, the pink particles either all 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. Which possibility will actually occur depends on the value of the fair coin d_i, that is, flipped at the depinking time $2iT = D_{j+1}$. It is easy to see that the coin is independent of $\{\widehat{\mathsf{ink}}_\ell\}_{\ell \le j}$, and this implies that both possibilities are equally likely. $\square$

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

LEMMA 7.1. *The sequence* $\{\widehat{\mathsf{ink}}_j\}_{j \ge 0}$ *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*

$$p(a, b) \equiv \frac{1}{2}\left(\mathbb{I}_{\{b = a + \Delta(a)\}} + \mathbb{I}_{\{b = a - \Delta(a)\}}\right) \qquad (a, b \in \{0, 1, 2, \ldots, |V| - k + 1\}). \tag{7.0.2}$$

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

$$\mathsf{Fill} \equiv \left\{ \lim_{j \to +\infty} \widehat{\mathsf{ink}}_j = |V| - k + 1 \right\} \tag{7.0.3}$$

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

REMARK 7.2. The event Fill corresponds to the number of red particles converging to $|V|-k+1$, that is, 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 1.1.

PROOF OF LEMMA 7.1. 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

$$\begin{aligned}
\mathbb{P}(\mathsf{Fill}) &= \mathbb{P}(\widehat{\mathsf{ink}}_\infty = |V|-k+1) \\
(\widehat{\mathsf{ink}}_\infty \in \{0,|V|-k+1\}) &= \frac{\mathbb{E}[\widehat{\mathsf{ink}}_\infty]}{|V|-k+1} \\
(\{\widehat{\mathsf{ink}}_j\}_{j\in\mathbb{N}} \text{ bounded, cf. Prop. 7.2}) &= \frac{\lim_{j\to+\infty}\mathbb{E}[\widehat{\mathsf{ink}}_j]}{|V|-k+1} \\
(\text{mart. property} + \widehat{\mathsf{ink}}_0 = 1) &= \frac{\mathbb{E}[\widehat{\mathsf{ink}}_0]}{|V|-k+1} = \frac{1}{|V|-k+1}. \qquad \square
\end{aligned}$$

We will need one final lemma before we proceed.

LEMMA 7.2. *For all* $\mathbf{b} \in (V)_{k-1}$ *and* $t \ge 0$,

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

PROOF. This follows from the previous lemma if we can show that Fill and $\mathbf{z}_t^C$ are independent. To see this, simply notice that Fill is entirely determined by the coin flips d_i performed at the depinking times, whereas the value of $\mathbf{z}_t^C$ does not at all depend on these coin flips. $\square$

REMARK 7.3. It transpires from the above that the chameleon process conditioned on Fill is the same as the unconditional process, *except that* the coin flips d_i performed at depinking times are biased. This remark will be useful in the proof of Lemma 6.2 in Section 9.3.

8. Convergence to stationarity in terms of ink. In this section we will prove Lemma 6.1, used in the proof of Lemma 1.2 (cf. Section 6.1), 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 5.1, whereby any $\mathbf{x} = (\mathbf{x}(1), \dots, \mathbf{x}(k)) \in (V)_k$ is written as a pair $\mathbf{x} = (\mathbf{z}, x)$ with $\mathbf{z} = (\mathbf{x}(1), \dots, \mathbf{x}(k-1))$ 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 $R = \{x\}$, $P = \varnothing$ and $W = V \setminus (\mathbf{O}(\mathbf{z}) \cup \{x\})$ and recall from Proposition 5.2 that the chameleon process $\{(\mathbf{z}_t^C, R_t^C, P_t^C, W_t^C)\}_{t \geq 0}$ satisfies

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

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 x_t^I conditionally on $\mathbf{z}_t^I$.

LEMMA 8.1. *Given $\mathbf{x} = (\mathbf{z}, x) \in (V)_k$, let $\tilde{\mathbf{x}}_t^I = (\mathbf{z}_t^I, \tilde{x}_t^I)$ where, conditionally on $\mathbf{z}_t^I$, $\tilde{x}_t^I$ is uniform over $V \setminus \mathbf{O}(\mathbf{z}_t^I)$. Then*

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

where Fill *is the event defined in Lemma 7.1 (see also Remark 7.2).*

PROOF. We have seen that $\mathbf{z}_t^I$ and $\mathbf{z}_t^C$ have the same distribution; cf. the proof of Proposition 5.2. We deduce that

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

where the last equality follows from Lemma 7.2. On the other hand, (8.1.1) implies

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

We deduce that

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

We now combine this with formula (2.2.3) for $d_{\mathrm{TV}}(\cdot, \cdot\cdot)$.

$$\begin{aligned}d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\tilde{\mathbf{x}}_t^I]) &\leq \sum_{\mathbf{b} = (\mathbf{c}, b) \in (V)_k} \mathbb{E}\big[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}}\big(1 - \mathsf{ink}_t^{\mathbf{x}}(b)\big)\big] \\ &= \sum_{\mathbf{c} \in (V)_{k-1}} \mathbb{E}\left[\mathbb{I}_{\{\mathbf{z}_t^C = \mathbf{c}\} \cap \mathsf{Fill}} \sum_{b \in V \setminus \mathbf{O}(\mathbf{c})} \big(1 - \mathsf{ink}_t^{\mathbf{x}}(b)\big)\right]\end{aligned}$$

$$\begin{aligned}
[\text{sum over } b + (6.1.2)] &= \sum_{\mathbf{c}\in(V)_{k-1}} \mathbb{E}[\mathbb{I}_{\{\mathbf{z}_t^C=\mathbf{c}\}\cap\mathsf{Fill}}(|V|-k+1-\mathsf{ink}_t^{\mathbf{x}})] \\
(\text{sum over } \mathbf{c}) &= \mathbb{E}[\mathbb{I}_{\mathsf{Fill}}(|V|-k+1-\mathsf{ink}_t^{\mathbf{x}})] \\
(\text{apply Lemma 7.1}) &= 1 - \frac{\mathbb{E}[\mathsf{ink}_t^{\mathbf{x}}|\mathsf{Fill}]}{|V|-k+1}. \qquad \square
\end{aligned}$$

8.2. *Distance to the stationary distribution in terms of ink.*

PROOF OF LEMMA 6.1. We will prove the following stronger inequality:

$$\sup_{\mathbf{x},\mathbf{y}\in(V)_k} d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\mathbf{y}_t^I]) \le 2k \sup_{\mathbf{w}\in(V)_k} \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{w}}}{|V|-k+1}\middle|\mathsf{Fill}\right], \tag{8.2.1}$$

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: that is, there exists an $i \in [k]$ with $\mathbf{u}(i) \neq \mathbf{v}(i)$ and $\mathbf{u}(r) = \mathbf{v}(r)$ for $r \in [k] \setminus \{i\}$. 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 kth coordinate. Using the notation from Section 5.4, we write $\mathbf{x} = (\mathbf{z}, x)$ and $\mathbf{y} = (\mathbf{z}, y)$ for $\mathbf{z} \in (V)_{k-1}$ and $x \in V \setminus \mathbf{O}(\mathbf{z})$. Defining $\tilde{\mathbf{x}}_t^I = (\mathbf{z}_t^I, \tilde{x}_t^I)$ as in Section 8.1 and $\tilde{\mathbf{y}}_t^I$ similarly, we see that $\mathcal{L}[\tilde{\mathbf{x}}_t^I] = \mathcal{L}[\tilde{\mathbf{y}}_t^I]$ for all $t \ge 0$. We deduce that

$$\begin{aligned}
d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\mathbf{y}_t^I]) &\le d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\tilde{\mathbf{x}}_t^I]) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_t^I], \mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
&\quad + d_{\mathrm{TV}}(\mathcal{L}[\tilde{\mathbf{x}}_t^I], \mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
(\text{3rd. term} = 0) &= d_{\mathrm{TV}}(\mathcal{L}[\mathbf{x}_t^I], \mathcal{L}[\tilde{\mathbf{x}}_t^I]) + d_{\mathrm{TV}}(\mathcal{L}[\mathbf{y}_t^I], \mathcal{L}[\tilde{\mathbf{y}}_t^I]) \\
(\text{use Lemma 8.1}) &\le \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{x}}}{|V|-k+1}\middle|\mathsf{Fill}\right] \\
&\quad + \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{y}}}{|V|-k+1}\middle|\mathsf{Fill}\right] \\
&\le 2 \sup_{\mathbf{w}\in(V)_k} \mathbb{E}\left[1 - \frac{\mathsf{ink}_t^{\mathbf{w}}}{|V|-k+1}\middle|\mathsf{Fill}\right].
\end{aligned} \tag{8.2.2}$$

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 $r \le 2k$ 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

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

Applying (8.2.3) 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. We first show that, in the first two phases of the chameleon process—a constant color and a color-changing phase—, the number of red particles decreases in expectation by a constant factor.

LEMMA 9.1 (Proven in Section 9.1). *Consider a modified chameleon process where one drops condition* (2) *for a pinkening step*; *cf. Box* 5.1. *Assume also that* $k \le |V|/2$, $|V| \ge 300$ *and that the initial state* $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$ *with* $|P| < |R| \le |W|$. *If the phase length parameter* T *satisfies*

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

then

$$\mathbb{E}[|R^C_{2T\_}|] \le (1-c)|R|,$$

where $c = 1/1000 > 0$.

With this, we will show that the first depinking time has an exponential moment.

LEMMA 9.2 (Proven in Section 9.2). *Consider a chameleon process* (*without the modification in the previous lemma*) *with* $|V| \ge 300$ *and* $k \le |V|/2$, *started from an initial state* $(\mathbf{z}, R, P, W) \in \mathcal{C}_k(V)$ *with* $|P| = \varnothing$. *There exists a universal constant* $K > 0$ *such that if the phase length parameter* T *satisfies* $T \ge 20\mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$, *the first depinking time* D_1 *of this process satisfies*

$$\mathbb{E}\big[e^{D_1/(KT)}\big] \le e.$$

In Section 9.3 we deduce Lemma 6.2 from Lemma 9.2.

9.1. *Loss of red particles in the two first phases.*

PROOF OF LEMMA 9.1. Note that there is no depinking at time $t = 0$, since there are less pink particles than white or red ones in the state $(\mathbf{z}, R, P, W)$. Finally, the conditions on P, W, R and k imply

$$3|W| \ge |R| + |P| + |W| = |V| - k + 1 \ge |V|/2 \quad \Rightarrow \quad |W| \ge |V|/6.$$

The interval $(0, T]$ is a constant-color phase where black, red and white particles are simply moved around. Lemma 5.2 shows 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 (Proven in Section 9.1.1). *For all* $(a,b) \in (V)_2$ *and* $S, L \subset V$ *with* $S \times L \subset (V)_2$, $|L| \geq |V|/12$,

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

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

REMARK 9.1. 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 $T = \Omega(\mathsf{T}_{\mathsf{IP}(2,G)})$. Morris's 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 or may not be a pinkening time, depending on whether pinkening condition (1) is satisfied. We will nevertheless consider the maps

(9.1.1) $\qquad \tilde{I}_t \equiv I_{(T,t]}, \qquad T \leq t < 2T$; cf. the definition in Section 5.1.

We emphasize that $\tilde{I}_t$ does *not* correspond directly to the evolution of the chameleon process in the time interval $(T, 2T)$. Propositions 3.3 and 5.1 imply:

PROPOSITION 9.2 (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 ϕ_a be the *first time* of the form τ_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 ϕ_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 1. *The number of pinkening steps performed in time interval* $(T, 2T)$ *is at least the number of* $b \in I(W)$ *such that* $F_a = b$ *for some* $a \in I(R)$.

PROOF. Let $b \in I(W)$. Given the rules for color-changing phases (cf. Box 5.1), the particle at that location will move in the time interval $(T, 2T)$ according to $\tilde{I}_t$ until the first time $t = \tau_m \in (T, 2T)$ such that $\tilde{I}_{t-}(b) \in e_m$ and the other endpoint of e_m is white (if such a time exists). Now if $a \in I(R)$ satisfies $F_a = b$ and $\tau_m = \phi_a$, we have $\{\tilde{I}_{t-}(b), a\} = e_m$ and a must still be red at time $(\phi_a)_-$, since it was not contained in an edge before in this phase. It follows that the particle started from b must become pink by time ϕ_a. $\square$

The claim implies

$$|R^C_{2T\_}| = |R| - \#\text{ of pinkening steps in } (T, 2T] \tag{9.1.2}$$

$$\leq |R| - \sum_{b \in I(W)} \mathbb{I}_{\bigcup_{a\in I(R)}\{F_a = b\}}. \tag{9.1.3}$$

The sum in the RHS satisfies

$$\begin{aligned}\sum_{b\in I(W)} \mathbb{I}_{\bigcup_{a\in I(R)}\{F_a=b\}} \geq & \sum_{a\in I(R), b\in I(W)} \mathbb{I}_{\{F_a=b\}} \\ & - \sum_{\{a,a'\}\subset I(R), b\in I(W)} \mathbb{I}_{\{F_a=b\}}\mathbb{I}_{\{F_{a'}=b\}},\end{aligned} \tag{9.1.4}$$

and we obtain

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

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 independent of I; cf. Proposition 9.2. We deduce

$$\begin{aligned}&\sum_{(a,b)\in(V)_2} \mathbb{P}\big(a \in I(R), b \in I(W), F_a = b\big) \\ &\qquad = \sum_{(a,b)\in(V)_2} \mathbb{P}\big(a \in I(R), b \in I(W)\big)\mathbb{P}(F_a = b) \\ &(\text{use Proposition 9.1}) \geq \frac{|R|}{13|V|} \sum_{(a,b)\in(V)_2} \mathbb{P}(F_a = b) \\ &\qquad = \frac{|R|}{13|V|} \sum_{a\in V} \mathbb{P}(F_a \neq *).\end{aligned} \tag{9.1.6}$$

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

$$\mathbb{P}(F_a \neq *) \geq 1 - \mathbb{P}\big(\tilde{I}_{2T}(a) = a\big) = \mathbb{P}(a^R_T \neq a),$$

where $\{a_t^R\}_{t\geq 0}$ is a realization of $\mathsf{RW}(G)$ started from a. By the contraction principle and Proposition 2.1,

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

which implies

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

We deduce from (9.1.6) that

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

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

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

We *claim* that:

CLAIM 2 (Proven in Section 9.1.2). *For all* $(a, a', b) \in (V)_3$,

$$\mathbb{P}(F_a = b, F_{a'} = b) \leq \mathbb{P}(F_a = a', F_{a'} = b) + \mathbb{P}(F_{a'} = a, F_a = b).$$

Summing up over b above gives at most $\mathbb{P}(F_a = a') + \mathbb{P}(F_{a'} = a)$ in the RHS. Therefore,

$$\begin{aligned}
&\sum_{\{a,a'\}\subset V, b\in V\setminus\{a,a'\}} \mathbb{P}\big(a, a' \in I(R)\big)\mathbb{P}(F_a = b, F_{a'} = b) \\
&\qquad \leq \sum_{\{a,a'\}\subset V} \mathbb{P}\big(a, a' \in I(R)\big)\big(\mathbb{P}(F_a = a') + \mathbb{P}(F_{a'} = a)\big) \\
&\qquad = \sum_{(a,a')\in (V)_2} \mathbb{P}\big(a \in I(R), a' \in I(R)\big)\mathbb{P}(F_a = a') \\
&\text{(apply Prop. 9.1)} = \left(\frac{|R|}{|V|} + 2^{-10}\right) \sum_{(a,a')\in (V)_2} \mathbb{P}\big(a \in I(R)\big)\mathbb{P}(F_a = a') \\
&\left(\bigcup_{a'}\{F_a = a'\} = \{F_a \neq *\}\right) = \left(\frac{|R|}{|V|} + 2^{-10}\right) \sum_{a\in V} \mathbb{P}\big(a \in I(R)\big)\mathbb{P}(F_a \neq *)
\end{aligned}$$

$$\begin{aligned}
\big(\mathbb{P}(F_a \neq *) \le 1\big) &\le \left(\frac{|R|}{|V|} + 2^{-10}\right) \sum_{a\in V} \mathbb{P}\big(a \in I(R)\big) \\
&= \left(\frac{|R|}{|V|} + 2^{-10}\right) \mathbb{E}[|I(R)|] \\
&= \left(\frac{|R|}{|V|} + 2^{-10}\right) |R| \qquad \text{since } I = I_{(0,T]} \text{ is a bijection.}
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}[|R^C_{2T}| - |R|] &\le |R|\left(\frac{|R|}{|V|} + 2^{-10} - \frac{1}{14}\right) \\
&\le -|R|/30 \qquad \text{if } |R| \le |V|/28.
\end{aligned} \tag{9.1.8}$$

If $|R| > |V|/28$, we can still find a subset $R_0 \subset R$ of size $|R_0| = \lfloor |V|/28 \rfloor$. Since

$$\sum_{b\in I(W)} \mathbb{I}_{\{\exists a\in I(R): F_a = b\}} \ge \sum_{b\in I(W)} \mathbb{I}_{\{\exists a\in I(R_0): F_a = b\}}, \tag{9.1.9}$$

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

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

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

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

since $|R| \le |V|$. We deduce that

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

which gives the lemma together with (9.1.8). □

9.1.1. *Proof of the required estimates for the I map* (*Proposition* 9.1).

PROOF OF PROPOSITION 9.1. Recall that $T \ge 20\mathsf{T}_{\mathsf{IP}(2,G)}(1/4)$, therefore $T \ge 2\mathsf{T}_{\mathsf{IP}(2,G)}(2^{-10})$ by Proposition 2.1. By the contraction principle [1], this also implies that $T \ge \mathsf{T}_{\mathsf{RW}(G)}(2^{-10})$.

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

$$\begin{aligned}
\mathbb{P}\big((a,b) \in I(S) \times I(S)\big) &= \mathbb{P}(\{a,b\} \subset S^I_T) \\
(\text{negative correlation, Lemma 3.1}) &\le \mathbb{P}(a \in S^I_T)\mathbb{P}(b \in S^I_T) \\
(\mathcal{L}[I] = \mathcal{L}[I^{-1}], \text{Proposition 3.2}) &= \mathbb{P}\big(a \in I(S)\big)\mathbb{P}(b^I_T \in S) \\
\big(T \ge \mathsf{T}_{\mathsf{RW}(G)}(2^{-10})\big) &\le \mathbb{P}\big(a \in I(S)\big)\left(\frac{|S|}{|V|} + 2^{-10}\right).
\end{aligned}$$

As for the other inequality in the proposition, we have

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

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 2.2. We note that $|L|/|V|\geq(1/12)$ and $1-2^{-8}\geq 12/13$ to finish the proof. $\square$

9.1.2. *Proof of claim on F_a (Claim 2).*

PROOF OF CLAIM 2. It suffices to show that for $(a,a',b)\in(V)_3$,

$$\mathbb{P}(F_a=b, F_{a'}=b, \phi_a\leq\phi_{a'}) = \mathbb{P}(F_a=b, F_{a'}=a, \phi_a\leq\phi_{a'}). \tag{9.1.10}$$

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

$$(\mathcal{P}, \{e_n\}_n, \{c_n\}_n, \{d_i\}_i),$$

that maps L_b into R_b and vice-versa. We describe Φ in words: all values of d_i, $T<\tau_j\leq 2T$ 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 Φ has the desired properties. Φ is clearly measure-preserving, since ϕ_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 flipped. 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 $\hat{F}_a=F_a$. The two processes coincide for any time $T<t<\phi_a$. If L_b holds, we have $\phi_a=\tau_j<2T$ for some j, and 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 $c\in V\setminus\{a,F_a\}$.

$$\text{Under } L_b, \forall t\in[\phi_a, 2T]: \qquad (\hat{I}_t(a), \hat{I}_t(b)) = (\tilde{I}_t(b), \tilde{I}_t(a)).$$

Under L_b, the edge e_ℓ corresponding to $\tau_\ell=\phi_{a'}$ was of the form $e_\ell=\{a', \tilde{I}_{\phi_{a'}-}(b)\}$. This implies $e_\ell=\{a', \hat{I}_{\phi_{a'}-}(a)\}$ in the $\hat{I}_t$ process, and the latter must be in the event R_b. This shows that $\mathbb{P}(L_b)\leq\mathbb{P}(R_b)$. The opposite inequality follows from reversing roles of the two processes. $\square$

9.2. *Estimate for the first depinking time* (*Lemma* 9.2).

PROOF OF LEMMA 9.2. As in Lemma 9.1 we drop condition (2) for a depinking time, and notice that this change does *not* change the value (or the distribution) of D_1. The modification to the process also does not affect the end result of Lemma 5.1: 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.

Moreover, we assume without loss of generality that $|R| \leq |W|$, which implies that $|R_t^C| \leq |W_t^C|$ for $t < D_1$. This implies $|W_t^C| \geq |V|/6$ unless there are more pink particles than red ones at time $t < D_1$; this follows from the reasoning in the beginning of the proof of Lemma 9.1 in Section 9.1. Recall that each pinkening step remores a red particle and creates two pink ones. It follows that $|R_{2iT_-}^C| < 2|R|/3$ implies $D_1 \leq 2iT$, and

$$
\begin{aligned}
\forall i \in \mathbb{N}_+ \qquad \mathbb{P}(\hat{D}_1 > i) &\leq \mathbb{P}(|\hat{R}_i| \geq 2|R|/3, \hat{D}_1 > i-1) \\
&\leq \frac{3\mathbb{E}[\hat{R}_i \mathbb{I}_{\{\hat{D}_1 > i\}}]}{2|R|} \qquad (9.2.1)\\
&= \frac{3\mathbb{E}[\mathbb{E}[\hat{R}_i | \hat{\mathcal{F}}_{i-1}] \mathbb{I}_{\{\hat{D}_1 > (i-1)\}}]}{2|R|}, \qquad (9.2.2)
\end{aligned}
$$

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

We now estimate the integrand in (9.2.2). Lemma 5.1 and its proof impliy that

$$\mathbb{E}[|\hat{R}_i| | \hat{\mathcal{F}}_{i-1}]$$

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

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

By Lemma 9.1, we can ensure that

$$\mathbb{E}[|\hat{R}_i| | \hat{\mathcal{F}}_{i-1}] \leq (1-c)|\hat{R}_{i-1}| \qquad \text{if } |\hat{W}_{i-1}| \geq |V|/6 \quad \text{and} \quad |\hat{P}_{i-1}| < |\hat{R}_{i-1}|.$$

As noted before, these conditions are always satisfied in the event $\{\hat{D}_1 > (i-1)\}$, because there are less pink than red particles. We deduce

$$
\begin{aligned}
\forall i \in \mathbb{N}_+ \qquad \frac{\mathbb{E}[|\hat{R}_i| \mathbb{I}_{\{\hat{D}_1 > i\}}]}{|R|} &\leq \frac{\mathbb{E}[\mathbb{E}[|\hat{R}_i| | \hat{\mathcal{F}}_{i-1}] \mathbb{I}_{\{\hat{D}_1 > (i-1)\}}]}{|R|} \\
&\leq (1-c) \left\{ \frac{\mathbb{E}[|\hat{R}_{i-1}| \mathbb{I}_{\{\hat{D}_1 > i-1\}}]}{|R|} \right\} \\
(\ldots \text{induction} \ldots) &\leq (1-c)^i.
\end{aligned}
$$

This implies

$$\mathbb{P}(D_1 > 2iT) = \mathbb{P}(\hat{D}_1 > i) \le \frac{3(1-c)^i}{2}, \qquad c = 1/1000 \text{ universal.}$$

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

9.3. *Proof of Lemma* 6.2.

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

$$\mathbb{E}\big[e^{D_j(\mathbf{x})/(KT)}\big] \le e^j, \qquad K > 0 \text{ from Lemma 9.2;} \tag{9.3.1}$$

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

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

introduced in Lemma 5.1, henceforth called the hat process. This time we take the initial state

$$(\mathbf{z}, R, P, W) \equiv \big(\mathbf{z}, \{x\}, \varnothing, V \setminus \big(\mathbf{O}(\mathbf{z}) \cup \{x\}\big)\big)$$

corresponding to $\mathbf{x} = (\mathbf{z}, x)$ in the sense of Proposition 5.2. Also recall the definition $\hat{D}_i \equiv D_i(\mathbf{x})/2T$ and note that (9.3.1) is equivalent to

$$\mathbb{E}[e^{\hat{D}_j/K'}] \le e^j, \qquad K' = 2K. \tag{9.3.2}$$

This is valid for $j = 1$ due to Lemma 9.2. For $j > 1$, we recall the definition of the σ-fields $\hat{\mathcal{F}}_i$, recall that $\hat{D}_{j-1}$ is a stopping time for the hat process (cf. Lemma 5.1) and obtain

$$\mathbb{E}[e^{\hat{D}_j/K'}] \le \mathbb{E}\big[e^{\hat{D}_{j-1}/K'}\mathbb{E}\big[e^{(\hat{D}_j-\hat{D}_{j-1})/K'}|\hat{\mathcal{F}}_{\hat{D}_{j-1}}\big]\big]. \tag{9.3.3}$$

We will apply the strong Markov property of the hat process (cf. Lemma 5.1 again) to bound the conditional expectation in the RHS. The conditional law of $\hat{D}_j - \hat{D}_{j-1}$ 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 $\hat{P}_{\hat{D}_{j-1}} = P^C_{D_{j-1\_}} \neq \varnothing$; in fact, since depinking occurs at time D_{j-1}, we know that $|P^C_{D_{j-1\_}}| \ge \min\{|R^C_{D_{j-1\_}}|, |W^C_{D_{j-1\_}}|\}$. 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 9.2 that

$$\mathbb{E}\big[e^{(\hat{D}_j-\hat{D}_{j-1})/K'}|\hat{\mathcal{F}}_{\hat{D}_{j-1}}\big] \le e \qquad \text{almost surely,}$$

so that

$$\mathbb{E}[e^{\hat{D}_j/K'}] \le \mathbb{E}[e^{\hat{D}_{j-1}/K'}]e \le e^j \qquad \text{by induction.} \tag{9.3.4}$$

This proves (9.3.3) and (9.3.1).

To prove the lemma, notice that conditioning on Fill simply biases the coin flips d_i performed at depinking times; cf. Remark 7.3. 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. 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 parameters can be bounded solely in terms of the constituent parts? Nonsymmetric 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. Can one do something less indirect over 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 $k = |V|$ particles.
- Recall the heuristic assumption in the Introduction: $\mathsf{T}_{\mathsf{EX}(k,G)}(\varepsilon) \le C_1 \times \mathsf{T}_{\mathsf{RW}(G)}(\varepsilon/k)$ 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 $\mathsf{T}_{\mathsf{IP}(|V|,G)}(\varepsilon) = C_1 \times \mathsf{T}_{\mathsf{RW}(G)}(\varepsilon/|V|)$? Could it even be possible that $\mathsf{T}_{\mathsf{IP}(|V|,G)}(\varepsilon) \le \mathsf{T}_{\mathsf{RW}(|V|,G)}(\varepsilon)$, that is, the interchange process mixes at least as fast as independent random walkers? This would give Aldous's (now proven) conjecture on the spectral gap as a corollary.

APPENDIX A: MIXING BOUNDS FOR $\mathsf{EX}(k, G)$ VIA CANONICAL PATHS

We use asymptotic notation below as shorthand; see, for example, [2] for precise definitions. Let $G = (V, E, \{w_e\}_{e\in E})$ be a weighted graph. 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. The state space of $\mathsf{EX}(k, G)$ has cardinality $\binom{|V|}{k} = 2^{\Theta(|V|)}$ if $k = \Theta(|V|)$. By the results of [9], if $\rho_{\mathsf{EX}(k,G)}$ is the log-Sobolev constant of $\mathsf{EX}(k, G)$, then

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

Log-Sobolev inequality for the Bernoulli–Laplace model. Consider the complete graph $K_{|V|}$ where each edge has weight $1/|V|$. $\mathsf{EX}(k, K_{|V|})$ is the so-called *Bernoulli–Laplace model* with k particles, whose log Sobolev constant is of the order $\Theta(\ln(|V|^2/k(|V|-k)))$. Notice that this is $\Theta(1)$ for $k = \Theta(|V|)$.

Comparison argument. Now consider a general weighted graph $G = (V, E, \{w_e\}_{e\in E})$. 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

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

It is shown in the proof of [8, Theorem 3.1] that this comparison constant for the Dirichlet forms of $\mathsf{RW}(G)$ can be "lifted" with no loss to $\mathsf{EX}(k, G)$. The comparison principle for the log Sobolev constant [9] implies $\rho_{\mathsf{EX}(k,G)} = \Omega(\phi^{-1}(G))$. We deduce

$$\mathsf{T}_{\mathsf{EX}(k,G)}(1/4) = O(\phi(G) \ln |V|) \qquad \text{if } k = \Theta(|V|). \tag{A.0.1}$$

It can be very hard to find good upper bounds on $\phi(G)$ in general, but the general lower bound we will present implies that

$$\phi(G) \geq \frac{2\overline{\mathrm{dist}^2}}{\overline{d}}, \tag{A.0.2}$$

where $\overline{\mathrm{dist}^2}$ is the average over all $(x, y) \in V^2$ 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

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

We note that this is a lower bound, which we do not know how to achieve in the examples in Table 1.

APPENDIX B: THE TRAJECTORY OF $\widehat{\mathsf{ink}}_j$ GIVEN Fill

We use the facts proven in Section 7 to derive the technical estimate in Proposition 6.1 in the proof of Lemma 1.2; cf. Section 6.1.

PROOF OF PROPOSITION 6.1. We take the notation in Section 7 for granted, but omit the superscript **w** in this proof. Our first goal will be to show that, conditionally on Fill, $\{\widehat{\mathsf{ink}}_j\}_{j\geq 0}$ is still a Markov chain. Repeating the steps of the proof of Lemma 7.1, we note that

$$\mathbb{P}(\mathsf{Fill}|(\widehat{\mathsf{ink}}_i)_{i\leq j}) = \frac{\mathbb{E}[\widehat{\mathsf{ink}}_\infty|(\widehat{\mathsf{ink}}_i)_{i\leq j}]}{|V|-k+1} = \frac{\widehat{\mathsf{ink}}_j}{|V|-k+1} = \mathbb{P}(\mathsf{Fill})\widehat{\mathsf{ink}}_j.$$

We deduce from Bayes's rule and the Markovian property that

$$\begin{aligned}\mathbb{P}\left(\bigcap_{i=1}^{j}\{\widehat{\mathsf{ink}}_i = a_i\}|\mathsf{Fill}\right) &= \mathbb{P}\left(\bigcap_{i=1}^{j}\{\widehat{\mathsf{ink}}_i = a_i\}\right)a_j \\ (\text{Markov property for } \widehat{\mathsf{ink}}_j) &= p(1,a_1)p(a_1,a_2)\cdots p(a_{j-1},a_j)a_j \\ &= q(1,a_1)\cdots q(a_{j-1},a_j),\end{aligned}$$

where

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

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

PROPOSITION B.1. *Conditionally on* 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* $\widehat{\mathsf{ink}}_0 = 1$.

For the remainder of the proof, we will use this proposition to bound $1 - \widehat{\mathsf{ink}}_\ell/(|V|-k+1)$. Actually, 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 Fill, $I_\ell > 0$ always, hence Z_ℓ is a.s. well defined for all ℓ. Moreover, one can check that $1 - I_\ell \leq Z_\ell$ always. Therefore the lemma will follow from the estimate

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

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

CLAIM 3.

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

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 $\widehat{\mathsf{ink}}_i$, and $Z_{\ell-1} = 0 \Rightarrow Z_\ell = 0$. We deduce

$$\mathbb{E}^{\mathsf{Fill}}[Z_\ell] = \mathbb{E}^{\mathsf{Fill}}\left[\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \,\middle|\, \widehat{\mathsf{ink}}_{\ell-1}\right] Z_{\ell-1}\mathbb{I}_{\{Z_{\ell-1}\neq 0\}}\right]. \tag{B.0.1}$$

We now bound the conditional expectation in the RHS. We may assume that $\widehat{\mathsf{ink}}_{\ell-1} = r$ with $0 < r < |V| - k + 1$ (otherwise $Z_{\ell-1} = 0$). Thus we wish to bound

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \,\middle|\, \widehat{\mathsf{ink}}_{\ell-1} = r\right], \qquad 1 \le r \le |V| - k.$$

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

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

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

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \,\middle|\, \widehat{\mathsf{ink}}_{\ell-1} = r\right] = \frac{1}{2}\left(\frac{f(r + \Delta(r)) + f(r - \Delta(r))}{f(r)}\right). \tag{B.0.2}$$

Recall the formula for $\Delta(r)$ (cf. Proposition 7.3),

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

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

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

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}} \,\middle|\, \widehat{\mathsf{ink}}_{\ell-1} = r\right] = \frac{1}{2}\left(\sqrt{1 - \frac{\Delta(r)}{r}} + \sqrt{1 + \frac{\Delta(r)}{r}}\right). \tag{B.0.3}$$

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

$$\mathbb{E}^{\mathsf{Fill}}\left[\frac{Z_\ell}{Z_{\ell-1}}\,\Big|\,\widehat{\mathsf{ink}}_{\ell-1}=r\right]=1-\frac{1}{8}\left(\frac{\Delta(r)}{r}\right)^2\le\frac{71}{72}. \tag{B.0.4}$$

Case 2: $(|V|-k+1)/2<r\le|V|-k$. In this case (B.0.3) holds with $r'=|V|-k+1-r$ replacing r. Similar calculations imply that the conditional expectation is also $\le 71/72$ in this case.

Thus we see that in both cases

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

Plugging this into (B.0.1) gives

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

which completes the proof. □

Acknowledgments. We thank Ton Dieker and Prasad Tetali for useful discussions on the exposition. We also thank an anonymous referee for a long list of typos in the previous version, as well as for numerous suggestions.

REFERENCES

[1] ALDOUS, D. and FILL, J. A. Reversible Markov Chains and random walks on graphs. Book draft. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

[2] ALON, N. and SPENCER, J. H. (2000). *The Probabilistic Method*, 2nd ed. Wiley, New York. MR1885388

[3] ANDJEL, E. D. (1988). A correlation inequality for the symmetric exclusion process. *Ann. Probab.* **16** 717–721. MR0929073

[4] BENJAMINI, I. and MOSSEL, E. (2003). On the mixing time of a simple random walk on the super critical percolation cluster. *Probab. Theory Related Fields* **125** 408–420. MR1967022

[5] CAPUTO, P. and FAGGIONATO, A. (2007). Isoperimetric inequalities and mixing time for a random walk on a random point process. *Ann. Appl. Probab.* **17** 1707–1744. MR2358639

[6] CAPUTO, P., LIGGETT, T. M. and RICHTHAMMER, T. (2010). Proof of Aldous' spectral gap conjecture. *J. Amer. Math. Soc.* **23** 831–851. MR2629990

[7] COOPER, C., FRIEZE, A. and RADZIK, T. (2009). Multiple random walks and interacting particle systems. In *Automata, Languages and Programming. Part II. Lecture Notes in Computer Science* **5556** 399–410. Springer, Berlin. MR2544812

[8] DIACONIS, P. and SALOFF-COSTE, L. (1993). Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** 696–730. MR1233621

[9] DIACONIS, P. and SALOFF-COSTE, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** 695–750. MR1410112

[10] DIEKER, A. B. (2010). Interlacings for random walks on weighted graphs and the interchange process. *SIAM J. Discrete Math.* **24** 191–206. MR2600660

[11] FOUNTOULAKIS, N. and REED, B. A. (2008). The evolution of the mixing rate of a simple random walk on the giant component of a random graph. *Random Structures Algorithms* **33** 68–86. MR2428978

[12] LEE, T.-Y. and YAU, H.-T. (1998). Logarithmic Sobolev inequality for some models of random walks. *Ann. Probab.* **26** 1855–1873. MR1675008

[13] LEVIN, D. A., PERES, Y. and WILMER, E. L. (2009). *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI. MR2466937

[14] LIGGETT, T. M. (1974). A characterization of the invariant measures for an infinite particle system with interactions. II. *Trans. Amer. Math. Soc.* **198** 201–213. MR0375531

[15] LIGGETT, T. M. (1985). *Interacting Particle Systems. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **276**. Springer, New York. MR0776231

[16] LIGGETT, T. M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **324**. Springer, Berlin. MR1717346

[17] MONTENEGRO, R. and TETALI, P. (2006). Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.* **1** x+121. MR2341319

[18] MORRIS, B. (2006). Spectral gap for the zero range process with constant rate. *Ann. Probab.* **34** 1645–1664. MR2271475

[19] MORRIS, B. (2006). The mixing time for simple exclusion. *Ann. Appl. Probab.* **16** 615–635. MR2244427

[20] MORRIS, B. (2009). Improved mixing time bounds for the Thorp shuffle and L-reversal chain. *Ann. Probab.* **37** 453–477. MR2510013

[21] MORRIS, B. and PERES, Y. (2005). Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields* **133** 245–266. MR2198701

[22] PETE, G. (2008). A note on percolation on $\mathbb{Z}^d$: Isoperimetric profile via exponential cluster repulsion. *Electron. Commun. Probab.* **13** 377–392. MR2415145

[23] YAU, H.-T. (1997). Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probab. Theory Related Fields* **109** 507–538. MR1483598

IMPA
ESTRADA DONA CASTORINA, 110
RIO DE JANEIRO, RJ 22460-320
BRAZIL
E-MAIL: rimfo@impa.br