The Annals of Probability
2020, Vol. 48, No. 6, 3077–3123
https://doi.org/10.1214/20-AOP1455

THE EXCLUSION PROCESS MIXES (ALMOST) FASTER THAN INDEPENDENT PARTICLES

BY JONATHAN HERMON[1] AND RICHARD PYMAR[2]

[1] *Department of Mathematics, University of British Columbia, jhermon@math.ubc.ca*
[2] *Department of Economics, Mathematics and Statistics, Birkbeck, University of London, r.pymar@bbk.ac.uk*

Oliveira conjectured that the order of the mixing time of the exclusion process with k-particles on an arbitrary n-vertex graph is at most that of the mixing-time of k independent particles. We verify this up to a constant factor for d-regular graphs when each edge rings at rate $1/d$ in various cases:

(1) when $d = \Omega(\log_{n/k} n)$,

(2) when gap := the spectral-gap of a single walk is $O(1/\log^4 n)$ and $k \geq n^{\Omega(1)}$,

(3) when $k \asymp n^a$ for some constant $0 < a < 1$.

In these cases, our analysis yields a probabilistic proof of a weaker version of Aldous' famous spectral-gap conjecture (resolved by Caputo et al.). We also prove a general bound of $O(\log n \log\log n/\text{gap})$, which is within a $\log\log n$ factor from Oliveira's conjecture when $k \geq n^{\Omega(1)}$. As applications, we get new mixing bounds:

(a) $O(\log n \log\log n)$ for expanders,

(b) order $d \log(dk)$ for the hypercube $\{0, 1\}^d$,

(c) order $(\text{Diameter})^2 \log k$ for vertex-transitive graphs of moderate growth and for supercritical percolation on a fixed dimensional torus.

1. Introduction. The *symmetric exclusion process* EX(k) on a finite, connected graph $G = (V, E)$ (with vertex set V and edge set E) is the following continuous-time Markov process. In a configuration, each vertex is occupied by either a black particle or a white particle (where particles of the same colour are indistinguishable), such that the total number of black particles is $k < |V| =: n$. For each edge e independently, at the times of a Poisson process of rate $r_e > 0$, switch the particles at the endpoints of e. In this work, we take G to be d-regular and set $r_e \equiv 1/d$. The *interchange process* IP(k) is similarly defined, apart from the fact that we label the black particles by the set $[k] := \{1, \dots, k\}$, so that they become distinguishable.

The exclusion process is among the most fundamental and well-studied processes in the literature on interacting particle systems [32, 33], with ties to card shuffling [26, 27, 42], statistical mechanics [7, 21, 41] and numerous other processes (see, e.g., [31], Chapter 23 and [32]). Apart from having a rich literature on the model on infinite graphs, such as the lattices $\mathbb{Z}^d$, the exclusion process on finite graphs has been one of the major examples driving quantitative study of finite Markov chains. Couplings and random walks collision [1, 39], comparison techniques [10] (see the discussion in [39], Appendix A) log-Sobolev inequalities [12, 29, 43], path coupling [15, 30, 31, 42] and variants of the evolving sets method [9, 37–39] have been applied to this process. Sharp results have been obtained for certain graphs including the complete graph [28, 29], the discrete tori $(\mathbb{Z}/L\mathbb{Z})^d$ [37], the path [27] (including the asymmetric case [23, 24]), the cycle [26] and a variety of random graphs [39]. Bounds on the mixing time of the related interchange process have also been obtained for various graphs [20].

Received September 2018; revised February 2020.
MSC2020 subject classifications. Primary 60J27, 60K35; secondary 82C22.
Key words and phrases. Exclusion process, mixing-time, chameleon process, particle system.

For a continuous-time Markov process Q we denote by $t_{\mathrm{mix}}^{Q}(\varepsilon)$ the total-variation ε-mixing time of Q (see, e.g., (16)). When $\varepsilon = 1/4$, we omit it from this notation. Oliveira [39] showed that for some absolute constant C, for general graphs and rates,

$$\forall \varepsilon \in (0,1), \quad \max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \le C t_{\mathrm{mix}}^{\mathrm{RW}(1)} \log(n/\varepsilon), \tag{1}$$

where $\mathrm{RW}(r)$ is the process of $r \in \{1, \dots, n\}$ independent continuous-time random walks on G, each having the same transition rates $(r_e : e \in E)$. It was left as an open problem to determine whether the following stronger relation holds:

$$\forall \varepsilon \in (0,1), \quad t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \le C t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon). \tag{2}$$

A heuristic reasoning for this conjecture is the fact that the exclusion process satisfies a strong negative dependency property called negative association [5], which in some sense is even stronger than independence (see Section 2.4). One of the motivations given in [39] for (1) is that it serves as a proxy for (2), on which it is commented that "if at all true, is well beyond the reach of present techniques". Part of the appeal of (2) is its connection to Aldous' spectral-gap conjecture, now resolved by Caputo, Liggett and Richthammer [8], which asserts that the spectral-gaps of processes $\mathrm{EX}(k)$, $\mathrm{IP}(r)$, $\mathrm{RW}(1)$ are the same for all $r \in [n]$ and $k \in [n-1]$. A further discussion of connections to this conjecture can be found in Section 1.5.

In this work, we consider the mixing time of $\mathrm{EX}(k)$ for general finite d-regular graphs with rates $r_e \equiv \frac{1}{d}$ and obtain bounds in terms of the spectral-profile and relaxation-time. We obtain a general upper bound which is within a $\log\log n$ factor of Oliveira's conjecture when $k = n^{\Omega(1)}$ and prove the conjecture in certain special cases for all k (which includes hypergraphs). Finally, we give lower bounds on the mixing time of $\mathrm{EX}(k)$ in terms of independent random walks.

Note that $\mathrm{EX}(k)$ is in one-to-one correspondence with $\mathrm{EX}(n-k)$, as we may consider the set of vacant (white) vertices instead of the occupied (black) ones. Hence we may assume throughout that $k \le n/2$.

1.1. *Our main general results.* We present various bounds on $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ and show how they relate to verifying (2) in general, and for specific graphs. The first result we present bounds $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ in terms of $t_{\mathrm{rel}} := \frac{1}{\mathrm{gap}}$ (the *relaxation-time*, which is the inverse of the *spectral-gap*, the smallest positive eigenvalue of $-\mathcal{L}$ for $\mathcal{L}$ the generator of $\mathrm{RW}(1)$) and a quantity related to the decay of the heat-kernel of a random walk, denoted P_t. Specifically, for each $\epsilon \in (0,1)$, let (recall $n := |V|$)

$$r_*(\epsilon) := \inf\left\{t : \max_{v \in V} P_t(v,v) - 1/n \le \frac{\epsilon}{(\log n)^2}\right\}. \tag{3}$$

THEOREM 1.1 (General mixing bound). *There exist universal constants $C_{1.1}, c_{1.1} > 0$ such that for every n-vertex d-regular graph G with rates $r_e \equiv \frac{1}{d}$ we have that*

$$\forall \varepsilon \in (0,1), \quad \max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \le C_{1.1}\big(t_{\mathrm{rel}} + r_*(c_{1.1})\big) \log(n/\varepsilon). \tag{4}$$

In particular, $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \lesssim t_{\mathrm{mix}}^{\mathrm{RW}(\lceil\sqrt{n}\rceil)}(\varepsilon) + r_*(c_{1.1}) \log(n/\varepsilon)$.

For expanders $t_{\mathrm{rel}} \asymp 1$, while it follows from the spectral decomposition that $r_*(\epsilon) \asymp_\epsilon \log\log n$. Hence we obtain the bound $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim \log n \log\log n$ for expanders (Oliveira's conjecture gives $\log n$). In fact, this is the only natural example we have where $r_*(\epsilon) \gg t_{\mathrm{rel}}$. In general (for n-vertex regular graphs) it can be shown (see (21) in Section 2) that

$$r_*(\epsilon) \lesssim_\epsilon (\log n)^4 \wedge t_{\mathrm{rel}} \log\log n \tag{5}$$

from which we verify (2) if $t_{\mathrm{rel}} = \Omega((\log n)^4)$ and $k = n^{\Omega(1)}$. Moreover, we establish that (2) holds in general up to a $\log\log n$ factor for $k = n^{\Omega(1)}$.

Next, we bound $t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon)$ in terms of $t_{\mathrm{sp}}(\varepsilon)$, the bound on the ε L_∞-mixing time obtained via the spectral profile; see (26) for a definition and t_{rel}.

THEOREM 1.2 (Mixing for sublinear number of particles). *For each $\delta \in (0,1)$, there exist universal constants $C_{1.2}(\delta), C'_{1.2}(\delta) > 0$ such that for every n-vertex d-regular graph G with rates $r_e \equiv \frac{1}{d}$ and all $k \le n^\delta$ we have that*

$$(6) \qquad \forall \varepsilon \in (0,1), \quad t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \le C_{1.2}(\delta) t_{\mathrm{sp}}\left(\frac{\varepsilon}{k}\right) \le C'_{1.2}(\delta)\left[t_{\mathrm{sp}}\left(\frac{1}{2}\right) + t_{\mathrm{rel}} \log(k/\varepsilon)\right].$$

In particular, if $k \in [n^\beta, n^{1-\beta}]$ for some $\beta \in (0, 1/2)$ then

$$(7) \qquad \forall \varepsilon \in (0,1), \quad t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \lesssim t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon).$$

This result is one of our principal improvements to the main result of Oliveira [39] as it gives refined bounds for the case $k = n^{o(1)}$. By applying this theorem, we verify (2) for all k under the condition $t_{\mathrm{sp}}(\frac{1}{2}) \lesssim t_{\mathrm{rel}}$; see Corollary 1.8. This condition holds for vertex-transitive graphs of moderate growth and for supercritical percolation on a fixed dimensional torus $(\mathbb{Z}/L\mathbb{Z})^d$ (see Secction 11.1). In these cases, we obtain $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp (\mathrm{diam}(G))^2 \log k$ uniformly in $k \le n/2$. Morris [37] obtained the same bound for $G = (\mathbb{Z}/L\mathbb{Z})^d$ and Oliveira proved the same bound on the giant component of supercritical percolation on $(\mathbb{Z}/L\mathbb{Z})^d$ for $k = n^{\Omega(1)}$.

We now explain how (7) follows from (6). If $n \ge k = n^{\Omega(1)}$, then from the definition of the spectral profile we have $t_{\mathrm{sp}}(\frac{\varepsilon}{k}) \asymp t_{\mathrm{rel}} \log(n/\varepsilon)$ for each $\varepsilon \in (0,1)$ (we remark that the upper bound here holds for all $k \le n$). Further, it can be shown (see (20) in Section 2) that for such k and ε, $t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon) \asymp t_{\mathrm{rel}} \log(n/\varepsilon)$, and so we verify (2) for $k \asymp n^\delta$ with $\delta \in (0,1)$. For expanders, we obtain $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim_\delta \log n$ for $k \le n^\delta$.

In the seminal work [42] where he invented the so-called Wilson method, Wilson proved that for the hypercube $\{\pm 1\}^d$ one has that $t_{\mathrm{mix}}^{\mathrm{EX}(2^{d-1})} \gtrsim d^2$ [42], p. 308. He conjectured that $t_{\mathrm{mix}}^{\mathrm{EX}(2^{d-1})} \asymp d^2$ (to be precise, one may interpret the last sentence in [42], Section 9.1, as saying that $t_{\mathrm{mix}}^{\mathrm{IP}(2^d)} \lesssim d^2$, which was verified by the first named author and Salez [18] after this paper appeared online). Using Theorem 1.2, we show that for the hypercube we have $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim d \log(dk)$ uniformly in $k \le 2^{d-1}$; see Section 11.2 (in fact, we treat general product graphs). We also obtain a lower bound of the same order. To the best of our knowledge, previously the best available upper bound for the hypercube was $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim d^2 \log d$ and for expanders was $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim (\log n)^2$, both due to Oliveira [39] (see (1)).

The last main bound on $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ is again just in terms of t_{rel}, but requires the degree to be growing sufficiently fast.

THEOREM 1.3 (Mixing for graphs of high degree). *There exist universal constants $C_{\mathrm{deg}}, C_{1.3} > 0$ such that for every n-vertex d-regular graph G with rates $r_e \equiv \frac{1}{d}$ if $d \ge C_{\mathrm{deg}} \log_{n/k} n$ then*

$$(8) \qquad \forall \varepsilon \in (0,1), \quad t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \le C_{1.3} t_{\mathrm{rel}} \log(n/\varepsilon).$$

This theorem verifies (2) for $k = n^{\Omega(1)}$ when $d = \Omega(\log_{n/k} n)$.

1.2. *Lower bounds*. We provide now a general lower bound on $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ in terms of t_{rel}. We remark that there are few known general lower bounds on $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ in the existing literature.

Theorem 1.4 shows that under a mild delocalization assumption regarding some eigenvector corresponding to the spectral-gap, one has that $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \gtrsim t_{\mathrm{mix}}^{\mathrm{RW}(k)}$ when $k = n^{\Omega(1)}$. Proposition 1.5 provides a general condition ensuring that such delocalization holds. Moreover, Corollary 1.8 provides a sufficient condition for $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp t_{\mathrm{mix}}^{\mathrm{RW}(k)}$ for all k.

To motivate our result, consider an n-vertex regular expander and attach a path of length $L := \lceil \log n \rceil$ to one of its vertices. We expect that in this case $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim t_{\mathrm{rel}} \log L$, and so $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \ll t_{\mathrm{rel}} \log k$ for $k = (\log n)^{\omega(1)}$. This demonstrates that in general we cannot expect $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \gtrsim t_{\mathrm{rel}} \log k$. We now give a sufficient condition for this to hold. Here, we make no assumptions on G nor on the rates $\mathbf{r} := (r_e : e \in E)$. Recall that $\mathcal{L}$ denotes the generator of RW(1) and let $\pi := \mathrm{Unif}(V)$ be the stationary distribution. For $f, g \in \mathbb{R}^V$, define $\|f\|_p^p := \mathbb{E}_\pi[|f|^p] = \sum_x \pi(x)|f(x)|^p$ for $p \in (0, \infty)$ and $\|f\|_\infty := \max_{v \in V} |f(v)|$.

THEOREM 1.4. *Let $\lambda > 0$ be an eigenvalue of $-\mathcal{L}$ and $f \neq 0$ a corresponding eigenfunction. If $\varepsilon, \delta \in (0, 1/4)$ and $k \le n/2$ are such that $\|f\|_1 \ge k^{-1/4+\delta}\|f\|_2$ and $4\delta \log k - \log(16/\varepsilon) \ge 0$ then*

$$t_{\mathrm{mix}}^{\mathrm{EX}(k)}(1-\varepsilon) \ge \frac{1}{2\lambda}\big(4\delta \log k - \log(16/\varepsilon)\big).$$

Note that in order to apply Theorem 1.4 it suffices to find one eigenfunction f satisfying $\frac{\|f\|_1}{\|f\|_2} \ge k^{-\frac{1}{5}}$. Denote the eigenvalues of $-\mathcal{L}$ by $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$. In practice, when applying Theorem 1.4 one should pick $\lambda = \lambda_2$. Observe that $\|f\|_2 \le \sqrt{n}\|f\|_1$ for all f (not necessarily an eigenfunction).

Proposition 1.5 below provides a general upper bound on $\|f\|_2/\|f\|_1$ for an eigenfunction f corresponding to an eigenvalue $\lambda > 0$ of $-\mathcal{L}$ in terms of λ/c_{LS}, where $c_{\mathrm{LS}} = c_{\mathrm{LS}}^{\mathrm{RW}(1)}$ is the log-Sobolev constant of the graph (defined in (23) of Section 2.2).

PROPOSITION 1.5. *For (nonzero) $f \in \mathbb{R}^V$ such that $\mathcal{L}f = -\lambda f$, we have*

$$\log\big(\|f\|_2/2\|f\|_1\big) \le \lambda/c_{\mathrm{LS}}. \tag{9}$$

It is natural to expect that $t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ is at least "weakly" monotone in k for $k \le n/2$. While this is immediate for $t_{\mathrm{mix}}^{\mathrm{IP}(k)}$, we do not know how to show this for the exclusion process.

CONJECTURE 1.6 (Weak monotonicity of the mixing time in the number of particles). *There exists an absolute constant $C > 0$ such that if $k_1 \le k_2 \le n/2$ then $t_{\mathrm{mix}}^{\mathrm{EX}(k_1)} \le C t_{\mathrm{mix}}^{\mathrm{EX}(k_2)}$.*

Embarrassingly, we can resolve only the case when $k_1 = 1$.

PROPOSITION 1.7. *There exists an absolute constant $c > 0$ such that $\min_{k \in [n-1]} t_{\mathrm{mix}}^{\mathrm{EX}(k)} \ge c t_{\mathrm{mix}}^{\mathrm{RW}(1)}$.*

We remark that in Proposition 1.7 we make no assumption on G nor on the rates.

1.3. *On the interchange process.* As is the case in [39], our arguments can be used to upper-bound IP(k) as long as $k \le (1-\theta)n$ for some constant $\theta \in (0,1)$ (in this case constants $C_{1.1}$, $C_{1.2}$ and $C_{1.4}$ will depend on θ).

In a recent work [2], Alon and Kozma showed that in the regular case with $r_e = 1/d$ the L_∞-mixing-time of IP(n) is $\lesssim t_{\mathrm{mix}}^{\mathrm{RW}(1)} \log n$ (their result is more general, but contains an additional multiplicative term, which need not be of order 1 for general rates or when the graph is not regular). This is obtained via a comparison argument which hinges on an elegant use of the octopus inequality of Caputo, Liggett and Richthammer [8].

1.4. *Extensions and further applications.* We present a couple of ways in which some of our assumptions can be relaxed; for further details, see the Supplementary Material [17], Appendix B.

- The assumption of regularity can be replaced with an assumption on neighbouring vertices having comparable degrees. In this case, the results of Theorems 1.1–1.3 still hold subject to a few modifications.
- The requirement $d \ge C_{\deg} \log_{n/k} n$ in (8) can be replaced (under some additional conditions) with the assumption that the ℓth neighbourhood of each vertex is at least of size $C_{\deg} \log_{n/k} n$ for some fixed ℓ.

A sequence of Markov chains is said to exhibit *precutoff* if for some $\delta_n = o(1)$ the $1-\delta_n$ and the δ_n mixing times of the nth chain in the sequence are comparable, that is, $t_{\mathrm{mix}}^{(n)}(\delta_n) \asymp t_{\mathrm{mix}}^{(n)}(1-\delta_n)$. Our results imply pre-cutoff in various circumstances.

The following corollaries summarize various scenarios in which the bounds of Theorems 1.1–1.4 and Proposition 1.7 take particularly simple forms and pre-cutoff occurs. In each of the four following statements, we let $G_m = (V_m, E_m)$ be a sequence of finite d_m-regular graphs of increasing sizes n_m with rates $r_e^{(m)} \equiv \frac{1}{d_m}$. We emphasize the identity of the graph we are considering by adding it as a superscript or in parentheses.

COROLLARY 1.8 (Proof in [17], Appendix B.3). *If $t_{\mathrm{rel}}(G_m) \asymp t_{\mathrm{sp}}^{G_m}(\frac{1}{2})$, then uniformly in $k_m \le n_m/2$ we have*

$$t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \asymp t_{\mathrm{rel}}(G_m)\log(k_m+1) \asymp t_{\mathrm{mix}}^{\mathrm{RW}(k_m),G_m}. \tag{10}$$

Moreover, the sequence $(\mathrm{EX}(k_m), G_m)$ *exhibits a pre-cutoff, provided $k_m \gg 1$.*

COROLLARY 1.9 (Proof in [17], Appendix B.3). *If $t_{\mathrm{mix}}^{\mathrm{RW}(1),G_m} \asymp t_{\mathrm{sp}}^{G_m}(\frac{1}{2})$, then for all fixed* $\delta \in (0,1)$, *uniformly in $k_m \le n_m^{1-\delta}$ we have*

$$t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \asymp_\delta t_{\mathrm{mix}}^{\mathrm{RW}(1),G_m} + t_{\mathrm{rel}}(G_m)\log(k_m+1) \asymp t_{\mathrm{mix}}^{\mathrm{RW}(k_m),G_m}. \tag{11}$$

Moreover, the sequence $(\mathrm{EX}(k_m), G_m)$ exhibits a pre-cutoff, provided that $t_{\mathrm{rel}}(G_m) \times \log(k_m+1) \gg t_{\mathrm{mix}}^{\mathrm{RW}(1),G_m}$ and $1 \ll k_m \le n_m^{1-\delta}$ for some $\delta \in (0,1)$.

COROLLARY 1.10 (Proof in [17], Appendix B.3). *There exist constants $c, c' > 0$ such that for all $\delta_m \in (0, 1/5)$ if $\frac{1}{c_{\mathrm{LS}}(G_m)} \le c\delta_m t_{\mathrm{rel}}(G_m)\log n_m$ for all m then*

$$t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \ge c'\big(t_{\mathrm{mix}}^{\mathrm{RW}(1),G_m} \vee \delta_m t_{\mathrm{rel}}(G_m)\log n_m\big) \tag{12}$$

for all m and all $k_m \in [n_m^{5\delta_m}, \frac{n_m}{2}]$.

COROLLARY 1.11 (Proof in [17], Appendix B.3). *Let* $\delta \in (0,1)$. *If* $\frac{1}{c_{\mathrm{LS}}(G_m)} \lesssim \frac{t_{\mathrm{rel}}(G_m)\log n_m}{\log\log n_m}$, *then*

$$t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \asymp t_{\mathrm{rel}}(G_m)\log(k_m+1) \asymp t_{\mathrm{mix}}^{\mathrm{RW}(k_m),G_m}, \tag{13}$$

uniformly for $k_m \in [n_m^\delta, \frac{n_m}{2}]$, *and the sequence* $(\mathrm{EX}(k_m), G_m)$ *exhibits a pre-cutoff provided* $n_m^\delta \le k_m \le n_m/2$.

If $\frac{1}{c_{\mathrm{LS}}(G_m)} \lesssim t_{\mathrm{rel}}(G_m)$ *then uniformly in* $k_m \le \frac{n_m}{2}$ *we have that*

$$t_{\mathrm{rel}}(G_m)\log(k_m+1) \lesssim t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \lesssim t_{\mathrm{rel}}(G_m)\log(k_m \vee \log n_m). \tag{14}$$

Moreover, if $k_m \gtrsim \log n_m$ *(and* $k_m \le n_m/2$*) the sequence* $(\mathrm{EX}(k_m), G_m)$ *exhibits a pre-cutoff and* $t_{\mathrm{mix}}^{\mathrm{EX}(k_m),G_m} \asymp t_{\mathrm{mix}}^{\mathrm{RW}(k_m),G_m}$.

1.5. *Aldous' spectral-gap conjecture.* In the spirit of Aldous' spectral-gap conjecture, now resolved by Caputo, Liggett and Richthammer [8], which asserts that the spectral-gaps of processes $\mathrm{EX}(k)$, $\mathrm{IP}(r)$, $\mathrm{RW}(1)$ are the same for all $r \in [n]$ and $k \in [n-1]$, one may conjecture the stronger relation

$$\begin{aligned} &\forall \mathbf{x} \in (V)_k, t \ge 0, \\ &\|\mathrm{P}_{\mathbf{x}}^{\mathrm{IP}(k)}(\mathbf{x}(t) \in \bullet) - \pi_{\mathrm{IP}(k)}\|_{\mathrm{TV}} \le \|\mathrm{P}_{\mathbf{x}}^{\mathrm{RW}(k)}(\mathbf{x}(t) \in \bullet) - \pi_{\mathrm{RW}(k)}\|_{\mathrm{TV}}. \end{aligned} \tag{15}$$

Observe that a positive answer to (15) will provide another proof to Aldous' conjecture. Indeed, (15) yields $t_{\mathrm{rel}}^{\mathrm{IP}(k)} \le t_{\mathrm{rel}}^{\mathrm{RW}(k)} = t_{\mathrm{rel}}^{\mathrm{RW}(1)}$, which can be deduced from (19). Conversely, the inequalities $t_{\mathrm{rel}}^{\mathrm{IP}(k)} \ge t_{\mathrm{rel}}^{\mathrm{EX}(k)} \vee t_{\mathrm{rel}}^{\mathrm{RW}(1)}$ for all $k \in [n]$ (where we define $t_{\mathrm{rel}}^{\mathrm{EX}(n)} = 0$) and $t_{\mathrm{rel}}^{\mathrm{EX}(k)} \ge t_{\mathrm{rel}}^{\mathrm{RW}(1)}$ for all $k \in [n-1]$ are the easier direction of Aldous' conjecture (see [8]). Similarly, our Theorems 1.1–1.3 show that for regular graphs $\max_k t_{\mathrm{rel}}^{\mathrm{EX}(k)} \lesssim t_{\mathrm{rel}} + r_*$ (recall that often $r_* \lesssim t_{\mathrm{rel}}$), while if $d \ge C_{\mathrm{deg}} \log_{k/n} n$ then $t_{\mathrm{rel}}^{\mathrm{EX}(k)} \lesssim t_{\mathrm{rel}}$, and (for all d) $\max_{k \le n^\delta} t_{\mathrm{rel}}^{\mathrm{EX}(k)} \lesssim_\delta t_{\mathrm{rel}}$. While this is of course weaker than the result of Caputo et al., what is interesting here is that our proof is entirely probabilistic.

It is plausible that Aldous' conjecture could be strengthened to an operator L_2 inequality, between the generator of the interchange process and that of the corresponding mean-field system (see [18], Conjecture 1). This would have striking consequences, including verifying Oliveira's conjecture (even for the L_2 mixing–time). See [2] for an application for the emergence of macroscopic cycles in the cycle decomposition of the permutation obtained by running the interchange process. We note that such an operator inequality was recently proved for the zero range process in [19].

QUESTION 1.12. Is it the case that there exists an absolute constant $C > 1$ and some nondecreasing continuous $f : [0,1] \to [0,1]$ with $f(0) = 0$ such that for all $t \ge 0$,

$$\begin{aligned} &\forall \mathbf{x} \in (V)_k, \\ &\quad \|\mathrm{P}_{\mathbf{x}}^{\mathrm{IP}(k)}(\mathbf{x}(t) \in \bullet) - \pi_{\mathrm{IP}(k)}\|_{\mathrm{TV}} \ge f(\|\mathrm{P}_{\mathbf{x}}^{\mathrm{RW}(k)}(\mathbf{x}(Ct) \in \bullet) - \pi_{\mathrm{RW}(k)}\|_{\mathrm{TV}}), \\ &\forall A \in \binom{V}{k}, k \le n/2 \\ &\quad \|\mathrm{P}_A^{\mathrm{EX}(k)}(A_t \in \bullet) - \pi_{\mathrm{EX}(k)}\|_{\mathrm{TV}} \ge f(\|\mathrm{P}_A^{\widehat{\mathrm{RW}}(k)}(\widehat{\mathbf{x}}(Ct) \in \bullet) - \pi_{\widehat{\mathrm{RW}}(k)}\|_{\mathrm{TV}}), \\ &\forall \varepsilon \in (0, 1/4), \\ &\quad C \max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) \ge t_{\mathrm{mix}}^{\mathrm{IP}(n)}(f(\varepsilon)), \end{aligned}$$

where $\widehat{\mathrm{RW}}(k)$ is the projection of $\mathrm{RW}(k)$ obtained by forgetting the labelling of the particles?

Asymptotic notation. We write $o(1)$ for terms which vanish as $n \to \infty$. We write $f_n = o(g_n)$ or $f_n \ll g_n$ if $f_n/g_n = o(1)$. We write $f_n = O(g_n)$ and $f_n \lesssim g_n$ (and also $g_n = \Omega(f_n)$ and $g_n \gtrsim f_n$) if there exists a constant $C > 0$ such that $|f_n| \le C|g_n|$ for all n. We write $f_n = \Theta(g_n)$ or $f_n \asymp g_n$ if $f_n = O(g_n)$ and $g_n = O(f_n)$. Throughout, $\log\log n$ is to be interpreted as $\log\log(n \vee e^e)$, where $a \vee b := \max\{a,b\}$ and $a \wedge b := \min\{a,b\}$.

Organization of the paper. In Section 2, we recall some properties of the exclusion process (its graphical construction and negative association), prove Proposition 1.5, show how the mixing time of k particles is related to the mixing time of one particle conditioned on the others and provide an auxiliary bound on the L_2 distance. In Section 3, we introduce the chameleon process as the main tool which allows us to bound the mixing time of one particle conditioned on the others. We also prove Theorem 1.1 subject to some technical propositions (the majority of whose proofs appear in the Supplementary Material [17]), the most significant of which being Proposition 3.3. We give a detailed overview of how we use the chameleon process in Section 4 and turn these heuristics into formal arguments in Section 5 and Section 6, proving Proposition 3.3 for the case of large degree. In Section 7, we show how to modify the arguments already presented in order to prove Proposition 3.3 for small degree graphs, as well as Theorems 1.2 and 1.3. We present the proof of the lower bounds in Section 10, and give further applications of our results in Section 11.

2. Preliminaries.

2.1. *Mixing times.* Note that since $\mathrm{EX}(k)$ and $\mathrm{IP}(k)$ are irreducible and have symmetric transition rates, the uniform distributions on their state spaces $\binom{V}{k}$ (the set of all subsets of V of size k) and $(V)_k$ (the set of all k-tuples of distinct vertices), respectively, are stationary. Recall that the total variation distance of two distributions on a finite set Ω is

$$\|\mu - \nu\|_{\mathrm{TV}} := \sum_{a:\mu(a)>\nu(a)} \big(\mu(a) - \nu(a)\big).$$

Throughout, we use the convention that $(X_t)_{t\ge0}$ is a continuous-time random walk on the graph G with the same jump rates as above (i.e., a realisation of $\mathrm{EX}(1)$), and that $(A_t)_{t\ge0}$ and $(\mathbf{x}(t))_{t\ge0}$ are $\mathrm{EX}(k)$ and $\mathrm{IP}(k)$, respectively (we sometimes use $(\mathbf{w}(t))_{t\ge0}$, $(\mathbf{y}(t))_{t\ge0}$ or $(\mathbf{z}(t))_{t\ge0}$ instead of $(\mathbf{x}(t))_{t\ge0}$). We denote the uniform distribution on V by π and on $\binom{V}{k}$ and $(V)_k$ by $\pi_{\mathrm{EX}(k)}$ and $\pi_{\mathrm{IP}(k)}$. We write P_x (resp., $\mathrm{P}_A^{\mathrm{EX}(k)}$, $\mathrm{P}_{\mathbf{x}}^{\mathrm{IP}(k)}$) for the law of $(X_t)_{t\ge0}$ given $X_0 = x$ (resp., $(A_t)_{t\ge0}$ given $A_0 = A$, $(\mathbf{x}(t))_{t\ge0}$ given $\mathbf{x}(0) = \mathbf{x}$). The total variation ε-mixing times of a single walk and of $\mathrm{EX}(k)$ are

$$t_{\mathrm{mix}}(\varepsilon) = t_{\mathrm{mix}}^{\mathrm{RW}(1)}(\varepsilon) := \inf\Big\{t : \max_{x\in V}\|\mathrm{P}_x(X_t \in \bullet) - \pi\|_{\mathrm{TV}} \le \varepsilon\Big\}, \tag{16}$$

$$t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\varepsilon) := \inf\Big\{t : \max_{A\in\binom{V}{k}}\|\mathrm{P}_A^{\mathrm{EX}(k)}(A_t \in \bullet) - \pi_{\mathrm{EX}(k)}\|_{\mathrm{TV}} \le \varepsilon\Big\}. \tag{17}$$

The mixing times $t_{\mathrm{mix}}^{\mathrm{IP}(k)}(\varepsilon)$ and $t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon)$ of $\mathrm{IP}(k)$ and $\mathrm{RW}(k)$, respectively, are analogously defined. Recall that when $\varepsilon = 1/4$ we omit it from the above notation.

The δ L_∞-mixing time of a single walk (throughout, we consider the L_2 and L_∞ distances and mixing times only w.r.t. a single walk) is defined as

$$t_{\mathrm{mix}}^{(\infty)}(\delta) := \inf\Big\{t : \max_{x,y\in V}\big|nP_t(x,y) - 1\big| \le \delta\Big\}$$

and we set $t_{\mathrm{mix}}^{(\infty)} := t_{\mathrm{mix}}^{(\infty)}(1/2)$. Recall that P_t denotes the heat-kernel of a single walk. The relaxation-time is defined as

$$t_{\mathrm{rel}} := \frac{1}{\mathrm{gap}} = \lim_{t\to\infty}\frac{-t}{\log[\max_{x\in V} P_t(x,x) - 1/n]},$$

that is, it is the inverse of the *spectral-gap*, the smallest positive eigenvalue of $-\mathcal{L}$, where $\mathcal{L}$ is the generator of a single walk.

We now note that we can characterize $t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon)$ in terms of $t_{\mathrm{mix}}^{\mathrm{RW}(1)}(\varepsilon/k)$, which in turn can be characterized in terms of the relaxation-time when $k = n^{\Omega(1)}$. This was used in the discussion following Theorem 1.2 (mixing for sublinear number of particles). Indeed,

$$\forall k \in \mathbb{N}, \varepsilon \in (0, 1/4), \quad \frac{1}{2} t_{\mathrm{mix}}^{\mathrm{RW}(1)}(4\varepsilon/k) \le t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon) \le t_{\mathrm{mix}}^{\mathrm{RW}(1)}(\varepsilon/k). \tag{18}$$

The second inequality is easy, while the first requires considering the separation distance, and noting that

$$\min_{\mathbf{x},\mathbf{y} \in V^k} \mathrm{P}_{\mathbf{x}}^{\mathrm{RW}(k)}(\mathbf{x}(t) = \mathbf{y}) = \Big[\min_{x,y \in V} \mathrm{P}_x^{\mathrm{RW}(1)}(X_t = y)\Big]^k;$$

cf. [25]. Generally, ([31], Lemma 20.6 and Lemma 20.11) for a Markov chain on a state space V of size n with a symmetric generator

$$\forall \varepsilon \in (0, 1), \quad t_{\mathrm{rel}} |\log \varepsilon| \le t_{\mathrm{mix}}^{\mathrm{RW}(1)}(\varepsilon/2) \le t_{\mathrm{mix}}^{(\infty)}(\varepsilon) \le t_{\mathrm{rel}} |\log n/\varepsilon|. \tag{19}$$

It follows by combining (18) and (19) that for all $C \ge 1$, $\varepsilon \in (0, 1)$ and all $k \in [4\varepsilon n^{1/C}, (n/\varepsilon)^C]$

$$\frac{1}{2C} t_{\mathrm{rel}} \log\bigl(n/(2\varepsilon)\bigr) \le t_{\mathrm{mix}}^{\mathrm{RW}(k)}(\varepsilon) \le (C+1) t_{\mathrm{rel}} \log(n/\varepsilon). \tag{20}$$

We verify now the claimed bound on $r_*(\epsilon)$ of (5).

For (n-vertex) regular graphs, $P_t(v, v) - \frac{1}{n} \lesssim (t + 1)^{-1/2}$ (e.g., [4, 35]) for all t. Hence $r_*(\epsilon) \le C(\epsilon)(\log n)^4$ for some constant C depending only on ϵ. As

$$\forall t \ge 0, i \in \mathbb{N}, \quad P_{it}(v, v) - \frac{1}{n} \ge \left(P_t(v, v) - \frac{1}{n}\right)^i$$

(which follows via the spectral decomposition), by (19) (used in the third inequality) we get that

$$\begin{aligned} r_*(\epsilon) &\lesssim t_{\mathrm{mix}}^{(\infty)}\left(\frac{\epsilon n}{(\log n)^2}\right) \lesssim_\epsilon (\log n)^4 \wedge t_{\mathrm{mix}}^{(\infty)} \frac{\log\log n}{\log n} \\ &\lesssim (\log n)^4 \wedge t_{\mathrm{rel}} \log\log n. \end{aligned} \tag{21}$$

2.2. *The spectral-profile, evolving sets and log-Sobolev.* As the generator $\mathcal{L}$ is symmetric, it is self-adjoint with respect to the inner product on $\mathbb{R}^V$ induced by π, given by $\langle f, g\rangle_\pi = \mathbb{E}_\pi[fg] := \sum_x \pi(x) f(x) g(x)$. Recall that the *spectral-gap* is gap $:= \lambda_2$ satisfies

$$\lambda_2 := \min\{\mathcal{E}(h, h)/\operatorname{Var}_\pi h : h \in \mathbb{R}^V \text{ is nonconstant}\}, \tag{22}$$

where $\mathcal{E}(f, f) := \langle -\mathcal{L} f, f\rangle_\pi = \frac{1}{2}\sum_{x,y} \pi(x)\mathcal{L}(x, y)(h(x) - h(y))^2$.

Recall also that the *log-Sobolev constant* is given by

$$c_{\mathrm{LS}} := \inf\left\{\frac{\mathcal{E}(h, h)}{\operatorname{Ent}_\pi h^2} : h^2 \in (0, \infty)^V\right\}, \tag{23}$$

where $\operatorname{Ent}_\pi f := \mathbb{E}_\pi[f \log(f/\|f\|_1)]$.

Denote $\Lambda(\varepsilon) := \min\{\mathcal{E}(h, h)/\operatorname{Var}_\pi h : h \in \mathbb{R}^V, \pi(\operatorname{supp}(h)) \le \varepsilon\}$, where $\operatorname{supp}(h) := \{x \in V : h(x) \ne 0\}$ is the *support* of h.

We now recall a couple of results from [14]. While some of the results below were originally stated in the case where $\mathcal{L}$ is of the form $K - I$, where I is the identity matrix and K is a transition matrix of a discrete-time Markov chain (possibly with nonzero diagonal entries), they hold for general $\mathcal{L}$, as we can always write $\mathcal{L} := \max_x |\mathcal{L}(x,x)|(K-I)$ for some transition matrix K (possibly with positive diagonal entries). (All the quantities considered below scale linearly in $\max_x |\mathcal{L}(x,x)|$.)

PROPOSITION 2.1 ([14] Lemma 4.2). *For all $\varepsilon \in (0,1)$,*

$$(1-\varepsilon)\Lambda(\varepsilon) \ge c_{\mathrm{LS}} \log(1/\varepsilon).$$

REMARK 2.2. It was shown in [16] that $17/c_{\mathrm{LS}} \le \max_{\varepsilon \le 1/2} \frac{\log(1/\varepsilon)}{\Lambda(\varepsilon)}$.

PROPOSITION 2.3 ([14] Lemma 2.1). *For any (nonzero) $u \in \mathbb{R}_+^V$, we have that*

$$\frac{\mathcal{E}(u,u)}{\mathrm{Var}_\pi u} \ge \frac{1}{2}\Lambda\big(4\|u\|_1^2/\mathrm{Var}_\pi u\big).$$

PROOF OF PROPOSITION 1.5. Let $f \in \mathbb{R}^V$ satisfy $-\mathcal{L}f = \lambda f$. We assume $\|f\|_2 \ge 2\|f\|_1$, as otherwise there is nothing to prove. By Propositions 2.1 and 2.3 we have that

$$\begin{aligned}\lambda &\ge \frac{\mathcal{E}(f,f)}{\mathrm{Var}_\pi f} \ge \frac{1}{2}\Lambda\big(4\|f\|_1^2/\mathrm{Var}_\pi f\big) \ge \frac{1}{2}\Lambda\big(4\|f\|_1^2/\|f\|_2^2\big)\\ &\ge c_{\mathrm{LS}} \log\big(\|f\|_2/2\|f\|_1\big). \qquad \square\end{aligned}$$

Recall that the L_p norm of a signed measure σ is

$$\|\sigma\|_{p,\pi} := \|\sigma/\pi\|_p, \quad \text{where } (\sigma/\pi)(x) = \sigma(x)/\pi(x).$$

In particular, for a distribution μ its L_2 distance from π satisfies

$$\|\mu - \pi\|_{2,\pi}^2 := \|\mu/\pi - 1\|_2^2 = \mathrm{Var}_\pi(\mu/\pi).$$

Let $\mu_t := \mathrm{P}_\mu^t$ and $u_t := \mu_t/\pi$. It is standard that $\frac{d}{dt}\mathrm{Var}_\pi(u_t) = -2\mathcal{E}(u_t,u_t)$ (e.g., [31], p. 284). By (22), $\mathcal{E}(u_t,u_t) \ge \lambda_2 \mathrm{Var}_\pi(u_t)$ from which it follows that $\frac{d}{dt}\mathrm{Var}_\pi(u_t) \le -2\lambda_2 \times \mathrm{Var}_\pi(u_t)$, and so by Grönwall's lemma

$$\|\mu_t - \pi\|_{2,\pi}^2 \le \|\mu - \pi\|_{2,\pi}^2 \exp(-2\lambda_2 t). \tag{24}$$

This is the well-known Poincaré inequality. The ε *L_p-mixing time* is defined as

$$t_{\mathrm{mix}}^{(p)}(\varepsilon) := \inf\Big\{t : \max_x \|\mathrm{P}_x^t - \pi\|_{p,\pi} \le \varepsilon\Big\}.$$

It is standard (e.g., [14] or [31], Proposition 4.15) that for reversible Markov chains, for all $x \in V$ and t we have

$$\max_{x,y}\left|\frac{P_t(x,y)}{\pi(y)} - 1\right| = \max_x \frac{P_t(x,x)}{\pi(x)} - 1 \quad \text{and} \quad \|\mathrm{P}_x^t - \pi\|_{2,\pi}^2 = \frac{P_{2t}(x,x)}{\pi(x)} - 1. \tag{25}$$

Thus $t_{\mathrm{mix}}^{(\infty)}(\varepsilon^2) = 2t_{\mathrm{mix}}^{(2)}(\varepsilon)$ for all $\varepsilon \le (\max_x \frac{1-\pi(x)}{\pi(x)})^{1/2}$. The *spectral-profile* [14] and *isoperimetric-profile/evolving-sets* [38] bounds on the ε L_∞ mixing time are respectively given by

$$\begin{aligned} t_{\mathrm{sp}}(\varepsilon) &:= \int_{4/n}^{4/\varepsilon} \frac{2d\delta}{\delta\Lambda(\delta)},\\ t_{\mathrm{evolving-sets}}(\varepsilon) &:= \max_x\big|\mathcal{L}(x,x)\big| \int_{4/n}^{4/\varepsilon\wedge 1/2} \frac{4d\delta}{\delta\Phi^2(\delta)} + t_{\mathrm{rel}}\log(8/\varepsilon)\mathbf{1}_{\{\varepsilon\le 8\}},\end{aligned} \tag{26}$$

where $\Phi(\delta) := \inf\{\frac{\sum_{a\in A, b\notin A}\pi(a)\mathcal{L}(a,b)}{\pi(A)} : A \subset V \text{ such that } \pi(A) \le \delta\}$. A generalization of the well-known discrete Cheeger inequality is that ([14], Lemma 2.4)

$$\Phi^2(\delta)/\Big(2\max_x\big|\mathcal{L}(x,x)\big|\Big) \le \Lambda(\delta) \le \Phi(\delta)/(1-\delta), \tag{27}$$

from which it follows that $t_{\mathrm{sp}}(\varepsilon) \le t_{\mathrm{evolving-sets}}(\varepsilon)$. Theorem 1.1 in [14] asserts that

$$\forall \varepsilon \in (0, n], \quad t_{\mathrm{mix}}^{(\infty)}(\varepsilon) \le t_{\mathrm{sp}}(\varepsilon) \le t_{\mathrm{evolving-sets}}(\varepsilon). \tag{28}$$

Plugging the estimate of Proposition 2.1 in (26) and then integrating over δ gives [14], Corollary 4.1 (cf. [22] for a slightly different argument).

PROPOSITION 2.4. *There exists an absolute constant C such that*

$$t_{\mathrm{sp}}\left(\frac{1}{2}\right) \le C\frac{\log\log n}{c_{\mathrm{LS}}}.$$

Using Proposition 2.3 (noting that $\|u_t\|_1 = 1$), the following refines (24).

PROPOSITION 2.5 ([14] Theorem 1.1). *For any initial distribution μ, we have that*

$$\|\mu_t - \pi\|_{2,\pi}^2 \le M, \quad \text{if } t \ge \int_{4/\|\mu-\pi\|_{2,\pi}^2}^{4/M} \frac{d\delta}{\delta\Lambda(\delta)}. \tag{29}$$

In particular, for all $0 < c < 1$ we have that

$$\|\mu_t - \pi\|_{2,\pi}^2 \le c\|\mu - \pi\|_{2,\pi}^2, \quad \text{if } t \ge \frac{\log(1/c)}{\Lambda(4/c\|\mu-\pi\|_{2,\pi}^2)}. \tag{30}$$

The following lemma is a simple consequence of Proposition 2.1 together with (29).

LEMMA 2.6. *Let r_* be as in (3). For every $c > 0$, we have that*

$$r_*(c) \lesssim_c \frac{\log\log n}{c_{\mathrm{LS}}\log n}. \tag{31}$$

2.3. *Graphical construction.* We present a graphical construction of the processes EX(k), IP(k) and RW(1), similar to that of Liggett [32] and Oliveira [39]. This construction enables us to define the processes on the same probability space, to then allow for direct comparison. We consider the following two ingredients:

1. a Poisson process Λ of rate $\frac{1}{d}|E|$;
2. an i.i.d. sequence of uniformly-distributed E-valued random variables $\{e_n\}_{n\in\mathbb{N}}$.

Next, we define the transpositions $f_e : V \to V$ for $e = \{u, v\} \in E$ as

$$f_e(x) = \begin{cases} u & \text{if } x = v, \\ v & \text{if } x = u, \\ x & \text{otherwise.} \end{cases}$$

We extend f_e to act on subsets of V and k-tuples by setting $f_e(A) = \{f_e(a) : a \in A\}$ and $f_e(\mathbf{x}) = (f_e(\mathbf{x}(1)), \dots, f_e(\mathbf{x}(k)))$. Then for $0 \le s \le t < \infty$ we define permutations $I_{[s,t]}$ as $I_{[s,t]} = f_{e_{\Lambda[0,t]}} \circ f_{e_{\Lambda[0,t]-1}} \circ \cdots \circ f_{e_{\Lambda[0,s)+1}}$, for $\Lambda[s,t] > 0$ (denoting the number of instances of the Poisson process Λ during time interval $[s,t]$), otherwise we set $I_{[s,t]}$ to be the identity map. Hence $I_{[s,t]}$ is the composition of the transpositions f_{e_j} that are chosen during $[s,t]$ composed in the order they occur. The following proposition is fundamental and its proof follows by inspection.

PROPOSITION 2.7 (Proof omitted). *Fix $t > 0$. Then*:

1. *For each $u \in V$, the process $\{I_{[s,s+t]}(u)\}_{t\geq 0}$ is a realisation of* RW(1) *initialised at u at time s.*
2. *For each $A \in \binom{V}{k}$, the process $\{I_{[s,s+t]}(A)\}_{t\geq 0}$ is a realisation of* EX(k) *initialised at A at time s.*
3. *For each $\mathbf{x} \in (V)_k$, the process $\{I_{[s,s+t]}(\mathbf{x})\}_{t\geq 0}$ is a realisation of* IP(k) *initialised at $\mathbf{x}$ at time s.*

2.4. *Negative association.* Let $Y_1, \dots, Y_m$ be real-valued random variables. Let $\mathbf{Y}_A := (Y_a)_{a\in A}$. We say that they are *negatively correlated* if $\mathrm{Cov}(Y_i, Y_j) \leq 0$ for all $i \neq j$. We say that they are *negatively associated* if

$$\text{(NA)} \qquad \mathbb{E}\big[f(\mathbf{Y}_A)g(\mathbf{Y}_B)\big] \leq \mathbb{E}f(\mathbf{Y}_A)\mathbb{E}g(\mathbf{Y}_B),$$

for all disjoint $A, B \subset [m]$ and all f, g nondecreasing w.r.t. the coordinatewise partial order $\leq^i_{\mathrm{cw}}$ on $\mathbb{R}^i$ (for $i = |A|, |B|$, resp.) defined via $(x_1, \dots, x_i) \leq^i_{\mathrm{cw}} (y_1, \dots, y_i)$ if $x_j \leq y_j$ for all $j \in [i]$. We say they are *conditionally negatively associated* (**CNA**) if for all $D \subset [m]$ the same holds when conditioning on $\mathbf{Y}_D$, that is,

$$\text{(CNA)} \qquad \begin{aligned} &\forall D \subset [m], \\ &\mathbb{E}\big[f(\mathbf{Y}_A)g(\mathbf{Y}_B) \mid \mathbf{Y}_D\big] \leq \mathbb{E}\big[f(\mathbf{Y}_A) \mid \mathbf{Y}_D\big]\mathbb{E}\big[g(\mathbf{Y}_B) \mid \mathbf{Y}_D\big] \end{aligned}$$

for all disjoint A, B and all nondecreasing f, g. Borcea, Brändén and Liggett [5] showed that (for the exclusion process) $(\mathbf{1}_{\{v\in A_t\}} : v \in V)$ is CNA, when A_0 is either deterministic or a product measure. It follows by taking the limit as $t \to \infty$ that the CNA property holds also for the stationary distribution $\pi_{\mathrm{EX}(k)} = \mathrm{Unif}(\binom{V}{k})$ (i.e., for $(\mathbf{1}_{\{v\in A\}} : v \in V)$, when $A \sim \pi_{\mathrm{EX}(k)}$).

It is clear that the NA property implies pairwise negative correlation (i.e., $\mathrm{Cov}(\mathbf{1}_{\{v\in A_t\}}, \mathbf{1}_{\{u\in A_t\}}) \leq 0$). While in [39] only the negative correlation property was used, we will make crucial use of the CNA property.

2.5. *From mixing of k particles to mixing of* 1 *particle conditioned on the rest.* By the contraction principle, it suffices to bound the mixing time of IP(k) as for all k,

$$\begin{aligned} &\max_{A\in\binom{V}{k}} \big\|\mathrm{P}^{\mathrm{EX}(k)}_A[A_t \in \bullet] - \pi_{\mathrm{EX}(k)}(\bullet)\big\|_{\mathrm{TV}} \\ (32)\qquad &\leq \max_{\mathbf{x}\in(V)_k} \big\|\mathrm{P}^{\mathrm{IP}(k)}_{\mathbf{x}}\big[\mathbf{x}(t) \in \bullet\big] - \pi_{\mathrm{IP}(k)}(\bullet)\big\|_{\mathrm{TV}} \leq \max_{\mathbf{x},\mathbf{y}\in(V)_k} \Delta_{\mathbf{x},\mathbf{y}}(t), \\ &\text{where } \Delta_{\mathbf{x},\mathbf{y}}(t) := \max_{\mathbf{x},\mathbf{y}\in(V)_k} \big\|\mathrm{P}^{\mathrm{IP}(k)}_{\mathbf{x}}\big[\mathbf{x}(t) \in \bullet\big] - \mathrm{P}^{\mathrm{IP}(k)}_{\mathbf{y}}\big[\mathbf{y}(t) \in \bullet\big]\big\|_{\mathrm{TV}}. \end{aligned}$$

We may interpolate between any two configurations $\mathbf{x}, \mathbf{y} \in (V)_k$ via a sequence of at most $k+1$ configurations, $\mathbf{x} = \mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_j = \mathbf{y} \in (V)_k$ such that $\mathbf{z}_i$ and $\mathbf{z}_{i-1}$ differ on exactly one coordinate for all $i \in [j]$. By symmetry, we may assume this is the kth coordinate (the total variation distance at time t w.r.t. two initial configurations is invariant under an application of the same permutation to their coordinates). By the triangle inequality, at a cost of picking up a factor k, we get that it suffices to consider two initial configurations which disagree only on their last coordinates:

$$\text{(33)} \qquad \max_{\mathbf{x},\mathbf{y}\in(V)_k} \Delta_{\mathbf{x},\mathbf{y}}(t) \leq k \max_{(\mathbf{w},y),(\mathbf{w},z)\in(V)_k:\mathbf{w}\in(V)_{k-1},y,z\in V} \Delta_{(\mathbf{w},y),(\mathbf{w},z)}(t).$$

Let $\mathbf{w}(t) = (\mathbf{w}_1(t), \dots, \mathbf{w}_{k-1}(t))$ be the positions of the first $k-1$ coordinates at time t. Given $\mathbf{w}(t)$, the positions of the kth coordinates at time t of both configurations on the right-hand side $y(t)$ and $z(t)$ converge (as $t \to \infty$) to the uniform distribution on $\mathbf{w}(t)^{\complement} := V \setminus \{\mathbf{w}_i(t) : i \in [k-1]\}$. It is thus natural to compare the two to $U \sim \mathrm{Unif}(\mathbf{w}(t)^{\complement})$ (given $\mathbf{w}(t)$) using the triangle inequality:

$$\begin{aligned}(34)\qquad &\max_{(\mathbf{w},y),(\mathbf{w},z)\in(V)_k:\mathbf{w}\in(V)_{k-1}} \Delta_{(\mathbf{w},y),(\mathbf{w},z)}(t) \\ &\le 2 \max_{(\mathbf{w},y)(V)_k:\mathbf{w}\in(V)_{k-1}} \|\mathcal{L}_{(\mathbf{w}(t),y(t))} - \mathcal{L}_{(\mathbf{w}(t),U)}\|_{\mathrm{TV}},\end{aligned}$$

where $\mathcal{L}_X$ denotes the law of X. Hence we reduced the problem of showing that $\Delta_{\mathbf{x},\mathbf{y}}(t) \le \varepsilon$ to that of showing that the maximum on the r.h.s. of (34) is at most $\frac{\varepsilon}{2k}$. The total-variation distance in the maximum is that of the last coordinate from $U \sim \mathrm{Unif}(\mathbf{w}(t)^{\complement})$, averaged over $\mathbf{w}(t)$. Hence loosely speaking, we reduced the problem to that of bounding the $\frac{\varepsilon}{2k}$-mixing time of the last coordinate, given the rest of the coordinates (in some averaged sense).

2.6. *An auxiliary lower bound on the L_2 distance.* Let $\mathscr{P}(V)$ be the collection of all distributions on V. For $A \subsetneq V$ and $\delta \in (0,1)$, let

$$\mathscr{P}_{A,\delta} := \{\mu \in \mathscr{P}(V) : \mu(A) \ge \pi(A) + \delta\pi(A^c)\}.$$

Note that $\nu_{A,\delta} := \delta\pi_A + (1-\delta)\pi \in \mathscr{P}_{A,\delta}$, where π_A denotes π conditioned on A (i.e., $\pi_A(a) = \pi(a)\mathbf{1}_{\{a\in A\}}/\pi(A)$). Moreover, $\min\{\delta' : \nu_{A,\delta'} \in \mathscr{P}_{A,\delta}\} = \delta$. It is thus intuitive that for a convex distance function between distributions, $\nu_{A,\delta}$ is the closest distribution to π in $\mathscr{P}_{A,\delta}$. The assertion of the following proposition can be verified using Lagrange multipliers, noting that the density function of the distribution with respect to π has to be constant on A and on $A^{\complement}$.

PROPOSITION 2.8 ([16] Proposition 4.1). *Let $A \subsetneq V$. Denote $\nu_{A,\delta} := \delta\pi_A + (1-\delta)\pi$. Then*

$$(35)\qquad \forall \delta \in (0,1) \qquad \min_{\mu\in\mathscr{P}_{A,\delta}} \|\mu - \pi\|_{2,\pi}^2 = \|\nu_{A,\delta} - \pi\|_{2,\pi}^2 = \delta^2\pi(A^{\complement})/\pi(A).$$

3. The chameleon process. Our main tool is the use of the *chameleon process*, a process invented by Morris [37] and used by Oliveira [39] and Connor-Pymar [9] to keep track of the distribution of a single particle in an interchange process, conditional on the locations of the other particles (see Proposition 3.4 for a precise formulation). As explained in Section 2.5, this can be used to upper bound the mixing time of the interchange process (and thus also of the exclusion process). This is quantified in Proposition 3.7. We will make use of several variants of this process. In some situations, the process consists of rounds of unvarying duration and is very similar to that used in [39]; whereas in others the length of rounds can vary in a way similar to [37]. The precise nature of the process depends on the values of k and d, and the current state of the process. We shall present first the version most similar to [39] (and with which we prove Theorems 1.1 (general mixing bound) and 1.3 (mixing for graphs of high degree)) and show in Section 9 how this can be adapted to prove Theorem 1.2 (mixing for sublinear number of particles).

3.1. *Description of the process.* We start this section with the construction of the chameleon process.

The first step is to modify slightly the graphical construction of Section 2.3. We suppose now that edges ring at rate $2/d$ and an independent fair coin flip determines whether particles on a ringing edge switch places or not. More formally, consider the following ingredients:

1. a Poisson process $\Lambda = \{\tau_1, \tau_2, \dots\}$ of rate $\frac{2}{d}|E|$;

2. an i.i.d. sequence of uniformly-distributed E-valued random variables $\{e_n\}_{n\in\mathbb{N}}$;
3. an i.i.d. sequence of coin flips $\{\theta_n\}_{n\in\mathbb{N}}$ with $\mathbb{P}(\theta_n = 1) = \mathbb{P}(\theta_n = 0) = 1/2$.

Recall the definition of f_e from Section 2.3 and set $f_e^1 = f_e$ and let f_e^0 be the identity function. We modify the definition of the maps $I_{[s,t]}$ from §2.3 as follows:

$$I_{[s,t]} = f_{e_{\Lambda[0,t]}}^{\theta_{\Lambda[0,t]}} \circ f_{e_{\Lambda[0,t]-1}}^{\theta_{\Lambda[0,t]-1}} \circ \cdots \circ f_{e_{\Lambda[0,s)+1}}^{\theta_{\Lambda[0,s)+1}}.$$

The joint distribution of the maps $I_{[s,t]}$, $0 \le s \le t < \infty$ is the same as in Section 2.3 by the thinning property of the Poisson process.

The choice of k in the following setup is relevant for obtaining an upper bound on $t_{\mathrm{mix}}^{\mathrm{IP}(k)}(\varepsilon)$. The chameleon process is a continuous-time Markov process built on top of the modified graphical construction and consisting of *burn-in periods*, and of *rounds*. We first describe a version in which the duration of each round is a fixed parameter t_{round}, known as the *round length* and to be chosen in the sequel. This version will be used to prove Theorems 1.1 (general mixing bound) and 1.3 (mixing for graphs of high degree). In the chameleon process, there is always one particle on each vertex, although not all particles are distinguishable. Each particle has an associated *colour*: one of black, red, pink and white. Formally, given a $(k-1)$-tuple $\mathbf{z} \in (V)_{k-1}$, let $\mathbf{O}(\mathbf{z}) := \{\mathbf{z}(1), \dots, \mathbf{z}(k-1)\}$ be the set of coordinates of $\mathbf{z}$. The state space of the chameleon process is given by

$$\Omega_k(V) := \{(\mathbf{z}, R, K, W) : \mathbf{z} \in (V)_{k-1}, \text{ and sets } \mathbf{O}(\mathbf{z}), R, K, W \text{ partition } V\}.$$

We denote the state at time t of the chameleon process started from $M_0 = (\mathbf{z}, R, K, W)$ as $M_t = (\mathbf{z}(t), \mathrm{R}_t, \mathrm{K}_t, \mathrm{W}_t)$. We say a particle at vertex v is *black* at time t if $v \in \mathbf{O}(\mathbf{z}(t))$, *red* if $v \in \mathrm{R}_t$, *pink* if $v \in \mathrm{K}_t$ and *white* if $v \in \mathrm{W}_t$. The black particles are distinguishable and their number remains constant throughout the process. We shall also denote the vector of positions of the black particles at time t by B_t (i.e., $\mathrm{B}_t = \mathbf{z}(t)$). By abuse of notation, we write $|\mathrm{B}_t|$ for $|\mathbf{O}(\mathbf{z}(t))|$, the number of black particles (note that B_t is a vector, not a set). Marginally, the evolution of B_t is simply that of the interchange process on $k-1$ particles, starting from $\mathbf{z}$. Conversely, the white (resp., pink and red) particles are indistinguishable, and their number changes as time varies. Suppose the chameleon process starts at time 0 from configuration $M_0 = (\mathbf{z}, R, \varnothing, W)$.

In order to define a quantity H_t we suppose that all particles are either unmarked or marked and at time t all particles are unmarked. Then suppose that at each instance during time interval $(t, t+1)$ at which an edge connecting an unmarked red particle and an unmarked white particle rings we mark both of these particles. We set H_t to be half the number of marked particles at time $t+1$.

We make the following definition.

DEFINITION 3.1. Let $\alpha \in (0, 1/4)$ and $t > 0$. We say that a configuration $M_0 = (\mathbf{z}, R, \varnothing, W)$ of the chameleon process is (α, t)-*good* if

$$\mathbb{E}_{M_0}[H_t] \ge 2\alpha(|R| \wedge |W|).$$

Let $p(M_0) = p(M_0, t) := \mathrm{P}_{M_0}[H_t \ge \alpha(|R| \wedge |W|)]$.

For an (α, t)-good configuration with $\alpha \le 1/4$, by Markov's inequality

$$\begin{aligned} p(M_0) &= 1 - \mathrm{P}_{M_0}[|R| \wedge |W| - H_t \ge (1-\alpha)(|R| \wedge |W|)] \\ &\ge 1 - \frac{\mathbb{E}_{M_0}[|R| \wedge |W| - H_t]}{(1-\alpha)(|R| \wedge |W|)} \ge \frac{\alpha}{1-\alpha} \ge \frac{4\alpha}{3}. \end{aligned} \tag{36}$$

Fix some $\alpha \in (0, 1/4)$ to be determined later. At time 0, we start with no pink particles. Similarly, at the beginning of each round we have that $\mathrm{K}_t = \varnothing$. We only start a round once we have an $(\alpha, t_{\mathrm{round}} - 1)$-good configuration. Initially, we let the process make successive burn-in periods, each of duration $t_{\mathrm{mix}}^{(\infty)}(n^{-10})$ and during which the process updates according to the updates of the underlying modified graphical construction, until the first time that at the end of a burn-in period we obtain an $(\alpha, t_{\mathrm{round}} - 1)$-good configuration. Similarly, if at the end of a round the configuration is not $(\alpha, t_{\mathrm{round}} - 1)$-good, then we let the process make successive burn-in periods, each of duration $t_{\mathrm{mix}}^{(\infty)}(n^{-10})$, until the first time that at the end of a burn-in period we obtain an $(\alpha, t_{\mathrm{round}} - 1)$-good configuration. Denote the beginning of the ith round by ρ_i and its end by $\hat{\tau}_i := \rho_i + t_{\mathrm{round}}$. We now describe a round of the chameleon process.

Each round consists of two phases. The first is a *constant-colour relaxation phase* of duration $t_{\mathrm{round}} - 1$, while the second is a *pinkening phase* of unit length. Loosely speaking, during a round the chameleon process evolves as the underlying interchange process, apart from the fact that pink particles are created by the recolouring of pairs of red and white particles (each pair consisting of a red and a white particle) during events known as *pinkenings*. Whenever an edge e_j rings at some time τ_j for which the two endpoints are occupied by a red and a white particle at this time, we colour both these particles pink, unless we have already obtained $2\lceil \alpha(|R| \wedge |W|)\rceil$ pink particles.

REMARK 3.2. One place in which our chameleon process differs from Oliveira's process is that we will always depink at the end of a round, whereas Oliveira waits to have a substantial number of pink particles before depinking.

The updates of the chameleon process during a single round are as follows:

- Intervals of time of the form $J_i := (\rho_i, \hat{\tau}_i - 1]$, for $i \in \mathbb{N}$, are *constant-colour phases* during which the chameleon process updates according to the updates of the underlying modified graphical construction, that is, if $t = \tau_j \in J_i$ for some $i \in \mathbb{N}$ then update as

$$\big(\mathbf{z}(t), \mathrm{R}_t, \varnothing, \mathrm{W}_t\big) = \big(f_{e_j}^{\theta_j}(\mathbf{z}(t_-)), f_{e_j}^{\theta_j}(\mathrm{R}_{t_-}), \varnothing, f_{e_j}^{\theta_j}(\mathrm{W}_{t_-})\big).$$

- Intervals of time of the form $\hat{J}_i := (\hat{\tau}_i - 1, \hat{\tau}_i)$, for $i \in \mathbb{N}$, are *pinkening phases* during which we update as in the constant-colour phase except for times $t = \tau_j \in \hat{J}_i$ at which both 1 and 2 below hold:

  1. e_j having a red endpoint $r \in \mathrm{R}_{t_-}$ and a white endpoint $w \in \mathrm{W}_{t_-}$,
  2. $|\mathrm{K}_{t_-}| < 2\lceil \alpha(|\mathrm{R}_{t_-}| \wedge |\mathrm{W}_{t_-}|)\rceil$.

  For such times, we update as

$$(\mathbf{z}_t, \mathrm{R}_t, \mathrm{K}_t, \mathrm{W}_t) = \big(\mathbf{z}_{t_-}, \mathrm{R}_{t_-} \setminus \{r\}, \mathrm{K}_{t_-} \cup \{r, w\}, \mathrm{W}_{t_-} \setminus \{w\}\big)$$

  and call t a *pinkening time*.
- Times of the form $t = \hat{\tau}_i$, for $i \in \mathbb{N}$, are called *depinking times* and are of two types:
  - Type 1 if $|\mathrm{K}_{t_-}| = 2\lceil \alpha(|\mathrm{R}_{t_-}| \wedge |\mathrm{W}_{t_-}|)\rceil$ and an independent biased coin $\hat{d}_i$ is equal to 1, where $\mathrm{P}[\hat{d}_i = 1 \mid M_{\rho_i}] = \frac{\alpha/2}{p(M_{\rho_i}, t_{\mathrm{round}}-1)}$ (recall that ρ_i is the beginning of the ith round). We then flip an independent fair (un-biased) coin d_i. If it lands heads ($d_i = 1$), we colour all pink particles red, and if it lands tails we colour all pink particles white.
  - Type 2 if $|\mathrm{K}_{t_-}| < 2\lceil \alpha(|\mathrm{R}_{t_-}| \wedge |\mathrm{W}_{t_-}|)\rceil$ or $\hat{d}_i = 0$. We then uniformly choose half of the pink particles (there is always an even number of pink particles) and colour these red, and the remaining half, we colour white.

Observe that as soon as $\mathrm{R}_t = \varnothing$ (resp., $\mathrm{W}_t = \varnothing$) it will remain empty while $|\mathrm{W}_s| = n - |\mathrm{B}_0|$ (resp., $|\mathrm{R}_s| = n - |\mathrm{B}_0|$) for all $s \geq t$. After such time, there will be no additional rounds.

Note that by (36) we have that $\mathrm{P}[\hat{d}_i = 1 \mid M_{\rho_i}] \leq 1$ and by definition of $p(\bullet, \bullet)$ we have that the probability of a type 1 depinking at time $\hat{\tau}_i$ is exactly $\alpha/2$ for all i (such that $|\mathrm{R}_{\rho_i}| \wedge |\mathrm{W}_{\rho_i}| \neq 0$). This means that if the number of red particles at the beginning of the round is r, then it stays r w.p. $1 - \alpha/2$, and otherwise with equal probability it changes to $r \pm \Delta(r)$, where $\Delta(r) := \lceil \alpha[r \wedge |\mathrm{W}_{\rho_i}|] \rceil = \lceil \alpha[r \wedge (n - |\mathrm{B}_0|)] \rceil$.

For $M_0 = (B, R, \varnothing, W)$ let $\hat{M}_t := (\hat{\mathrm{B}}_t, \hat{\mathrm{R}}_t, \hat{\mathrm{W}}_t)$ be the configuration at time t obtained from the modified graphical construction with $\hat{\mathrm{B}}_0 = B$, $\hat{\mathrm{R}}_0 = R$ and $\hat{\mathrm{W}}_0 = W$, that is, without any colour-changing of particles. The definition of (α, t)-good extends naturally to the process $\hat{M}_t$. Let $t_0 := t_{\mathrm{mix}}^{(\infty)}(n^{-10})$ and

$$\beta(\alpha, t) := \max_{B,R,W} \sup_{s \geq t_0} \mathrm{P}\big[\hat{M}_s \text{ is not } (\alpha, t)\text{-good} \mid \hat{M}_0 = (B, R, W)\big], \tag{37}$$

where the maximum is taken over all partitions of V into sets $\mathbf{O}(B)$, R, W with $B \in (V)_j$ for some $j \leq n/2$ satisfying $\{B(i) : i \in [j]\} = \mathbf{O}(B)$.

Recall the definition of $r_*(\epsilon)$ in (3). For each $\epsilon \in (0, 1)$, we define similar quantities:

$$\begin{aligned} t_*(\epsilon) &:= \inf\left\{t : \max_{v \in V} P_t(v, v) - 1/n \leq \frac{\epsilon}{\log n}\right\}, \\ s_*(\epsilon) &:= \inf\left\{t : \max_{v \in V} P_t(v, v) - 1/n \leq \frac{\epsilon}{t_*(\epsilon)}\right\}. \end{aligned} \tag{38}$$

For the proof of Theorem 1.1 (general mixing bound), we will show that for some positive constants α, C_{round}, ϵ, if we take

$$t_{\mathrm{round}} = C_{\mathrm{round}}\big(t_{\mathrm{rel}} + t_*(\epsilon) + s_*(\epsilon)\big) + 1$$

we have that $\beta(\alpha, t_{\mathrm{round}} - 1) \leq n^{-10}$. We state this as the following proposition.

PROPOSITION 3.3. *There exist constants $\epsilon, \alpha, C_{\mathrm{round}} > 0$, such that for all n sufficiently large*

$$\beta\big(\alpha, C_{\mathrm{round}}\big(t_{\mathrm{rel}} + t_*(\epsilon) + s_*(\epsilon)\big)\big) \leq n^{-10}.$$

We will explain in Section 3.3 how this implies the assertion of Theorem 1.1 (general mixing bound). For Theorem 1.3 (mixing for graphs of high degree), we show that it suffices to take $t_{\mathrm{round}} = C_{\mathrm{round}} t_{\mathrm{rel}} + 1$; see Proposition 8.1 (which is the analogue of the previous proposition). The situation is more involved for the proof of Theorem 1.2 (mixing for sublinear number of particles); see Section 9.

3.2. *Further technical results*. We present the key tools regarding the chameleon process that, together with Proposition 3.3, will be used to complete the proof of Theorem 1.1 (general mixing bound) in the following subsection.

Following Oliveira [39], we introduce a notion of *ink*, which represents the amount of *redness* either at a vertex or in the whole system. We write $\mathrm{ink}_t(v)$ for the amount of ink at vertex v at time t defined as $\mathrm{ink}_t(v) := \mathbf{1}_{\{v \in \mathrm{R}_t\}} + \frac{1}{2}\mathbf{1}_{\{v \in \mathrm{K}_t\}}$, and the amount of ink in the whole system at time t as $\mathrm{ink}_t := |\mathrm{R}_t| + \frac{1}{2}|\mathrm{K}_t|$. Notice that, by the construction of the chameleon process, the value of ink_t can only change at depinking times of Type 1. The following proposition links the amount of ink at a vertex to the probability that vertex is occupied by the kth particle, in a k-particle interchange process. The statement is identical to Proposition 5.2 of Oliveira (the difference being our chameleon process is constructed slightly differently). The proof is almost identical to the proof of Lemma 1 of [37], and we include our version for completeness.

PROPOSITION 3.4 (Proof in the Supplementary Material [17], Appendix C.1). *Consider a realisation $(\mathbf{x}(t))_{t\geq 0}$ of the k-particle interchange process started from configuration $\mathbf{x}=(\mathbf{z},x)$ and a corresponding chameleon process started from configuration $(\mathbf{z},\{x\},\varnothing,V\setminus(\mathbf{O}(\mathbf{z})\cup\{x\}))$. Then for each $t\geq 0$ and $\mathbf{b}=(\mathbf{c},b)\in(V)_k$, $\mathbf{c}\in(V)_{k-1}$,*

$$\mathrm{P}^{\mathrm{IP(k)}}\big[\mathbf{x}(t)=\mathbf{b}\big]=\mathbb{E}\big[\mathrm{ink}_t(b)\mathbf{1}_{\{\mathbf{z}(t)=\mathbf{c}\}}\big].$$

REMARK 3.5. Right after we colour two particles pink, since we do not reveal whether the edge ring of the edge connecting them was ignored or not, we cannot tell which one of them is at which location. The action of colouring them by pink symbolizes this uncertainty, which is the real reason that the assertion of the last proposition holds.

The next observation is that ink_t is a martingale. This can be readily checked from the behaviour of the chameleon process at depinking times. Moreover, as $t\to\infty$, ink_t converges to one of the two absorbing states 0 and $=n-k+1$. We define Fill as the event that this limit is $n-k+1$, that is, that eventually the only particles present in the system are red and black. One consequence of the martingale property of ink_t is that $\mathrm{P}[\mathrm{Fill}]=(n-k+1)^{-1}$.

LEMMA 3.6 (cf. [39] proof of Lemma 7.2). *The event* Fill *is independent of* $(\mathrm{B}_t:t\geq 0)$.

SKETCH PROOF. This follows from the fact that the coins $(d_i:i\in\mathbb{N})$ are independent of the coins $(\hat{d}_i:i\in\mathbb{N})$ and of the graphical representation. $\square$

Let us write $\widehat{\mathbb{E}}$ and $\widehat{\mathrm{P}}$ for the expectation and probability conditioned on the event Fill. We may add subscript $(\mathbf{w},y)\in(V)_k$ such that $\mathbf{w}\in(V)_{k-1}$ and $y\in V$ to indicate that the initial configuration of the interchange process is $(\mathbf{w},y)$, and thus for the chameleon process $\mathrm{R}_0=y$ and $\mathrm{B}_t=\mathbf{w}(t)$ for all t, where $\mathbf{w}(t)=(\mathbf{w}_1(t),\dots,\mathbf{w}_{k-1}(t))$ is the vector of the positions of the first $k-1$ coordinates at time t. In this case, we let $y(t)$ denote the position of the kth coordinate at time t. The main inequality relating the total-variation distance to the chameleon process is the following.

PROPOSITION 3.7 ([37], Lemma 2, [39], Lemma 6.1; proof in the Supplementary Material [17], Appendix C.2). *Let $\Delta_{\mathbf{x},\mathbf{y}}(t)$ be as in (32). Then*

$$\max_{\mathbf{x},\mathbf{y}\in(V)_k}\Delta_{\mathbf{x},\mathbf{y}}(t)\leq 2k\max_{(\mathbf{w},y)\in(V)_k:\mathbf{w}\in(V)_{k-1},y\in V}\widehat{\mathbb{E}}_{(\mathbf{w},y)}\big[1-\mathrm{ink}_t/(n-k+1)\big]. \tag{39}$$

The following proposition, which is essentially Proposition B.1 in [39], allows us to bound the right-hand side of (39). For $j\in\mathbb{N}$, we define event

$$A(j):=\big\{\text{config. at time } t(j) \text{ is not } (\alpha,t_{\mathrm{round}}-1)\text{-good}\big\}.$$

The term $t_{\mathrm{mix}}^{(\infty)}(n^{-10})$ below corresponds to the initial burn-in period, while the error term $\widehat{\mathrm{P}}_{(\mathbf{w},y)}[\bigcup_{j=0}^{i-1}A(j)]$ corresponds to the probability that additional burn-in periods occurred by the end of the ith round (i.e., that at time $t(j):=t_{\mathrm{mix}}^{(\infty)}(n^{-10})+jt_{\mathrm{round}}$ the configuration was not good). Hence, the assertion of the proposition is that the expected fraction of "missing ink" $1-\mathrm{ink}_t/(n-k+1)$ decays exponentially in the number of rounds.

PROPOSITION 3.8 (Proof in the Supplementary Material [17], Appendix C.3). *There exists* $c_\alpha \in (0,1)$ *such that for all* $i \in \mathbb{N}$ *and* $(\mathbf{w}, y) \in (V)_k$,

$$
\begin{aligned}
\widehat{\mathbb{E}}_{(\mathbf{w},y)}[1-\mathrm{ink}_{t(i)}/(n-k+1)] &\le \sqrt{n-k+1}\,c_\alpha^i + \widehat{\mathrm{P}}_{(\mathbf{w},y)}\left[\bigcup_{j=0}^{i-1} A(j)\right] \\
&\le \sqrt{n-k+1}\,c_\alpha^i + (\mathrm{P}[\mathrm{Fill}])^{-1}\mathrm{P}_{(\mathbf{w},y)}\left[\bigcup_{j=0}^{i-1} A(j)\right] \qquad (40)\\
&\le \sqrt{n-k+1}\,c_\alpha^i + i(n-k+1)\beta(\alpha, t_{\mathrm{round}}-1).
\end{aligned}
$$

3.3. *Proof of Theorem* 1.1 (*general mixing bound*).

PROOF. First, recall again that for n-vertex regular graphs, $P_t(v,v) - \frac{1}{n} \lesssim (t+1)^{-1/2}$ from which it follows that there exists a universal constant κ such that for any $\epsilon \in (0,1)$ $t_*(\epsilon) + s_*(\epsilon) \le 2r_*(\epsilon^3/\kappa^2)$. Next, using sub-multiplicativity [31], p. 54, we have that $t_{\mathrm{mix}}^{\mathrm{EX}(k)}((2n)^{-i}) \le i t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\frac{1}{4n})$. It follows that it suffices to consider $\varepsilon = \frac{1}{4n}$. We may assume n is at least some sufficiently large constant N (this was implicitly/explicitly used in several places), as there are only finitely many graphs for $n \le N$ (and hence finitely many processes, since we assume edge-rates are all $1/d$). Combining Propositions 3.3, 3.7 and 3.8 concludes the proof (use Proposition 3.8 with $i = \lceil \frac{4}{1-c_\alpha} \log n \rceil$, noting that the term $t_{\mathrm{mix}}^{(\infty)}(n^{-10})$ in the definition of $t(i)$ is $\lesssim t_{\mathrm{rel}} \log n$). □

4. An overview of our approach. The approach taken by Oliveira [39] is to let the constant-colour and the pinkening phases both be of order $t_{\mathrm{mix}}^{\mathrm{EX}(2)}$. The two main steps in his analysis are (i) to show that $t_{\mathrm{mix}} \asymp t_{\mathrm{mix}}^{\mathrm{EX}(2)}$ and (ii) by the choice of the duration of the constant-colour phase, using a delicate negative correlation argument deduce that with probability bounded from below a certain fraction of the red (or white, whichever set is of smaller size) particles will become pink in each pinkening phase. Both steps are much more difficult than what one might expect. As explained below, assuming regularity allows us to take our pinkening phase to be of duration of one time unit.

Since the red and the white particles play symmetric roles, we may assume that at the end of the last round prior to the current time we have $r \le (n-k+1)/2$ red particles (i.e., there are at least as many white particles as there are red; otherwise, switch their roles in what comes).

We now sketch the main ideas behind the proof of Proposition 3.3 in more detail. We will focus in this section and in Sections 5–6 on graphs with degree d satisfying $d \ge 10^4$, and describe how to extend the argument to all d in Section 7.2.

In order for a configuration to be (α, t)–good for some constant α, it suffices that (given the current configuration) with probability bounded from below, after t time units, at least some c-fraction of the red particles will have at least a c-fraction of their neighbours white. To see this, observe that if a red particle has $j \ge cd$ white neighbours, the chance an edge connecting it to any of them rings before the two particles at its end-point moved is $\frac{j}{2d-1} = \Omega(c)$ (and the probability this happens in at most 1 time unit is $\Omega(c)$).

Observe that if (at the end of a constant-colour phase) a vertex has at most $(\frac{r}{n} + \frac{c}{4})d$ red neighbours and at most $(\frac{k-1}{n} + \frac{c}{4})d$ black neighbours, then it has at least $\frac{1}{2}(1 - \frac{k-1}{n} - c)d \ge (\frac{1}{4} - \frac{c}{2})d$ white neighbours (as $r \le \frac{1}{2}(n-k+1)$). Hence, instead of controlling the number of white neighbours of a vertex, conditioned on it being red, we may control the number of red neighbours and the number of black neighbours separately. This is done is Section 5.1 and Section 5.2, respectively.

We say that two particles *interacted* if an edge connecting them rang. Exploiting the CNA property in conjunction with L_2-contraction considerations allows us to control the number of red with red interactions during the pinkening phase, provided that $t_{\mathrm{round}} - 1 \geq Ct_{\mathrm{rel}}$. This L_2 argument is the key that allows us to avoid taking $t_{\mathrm{round}} \geq Ct_{\mathrm{mix}}$, as Oliveira does.

Controlling the number of red with black interactions during a pinkening phase requires exploiting the NA property to derive certain large deviation estimates for the occupation measure of the black particles, as well as a certain decomposition which allows us to overcome the dependencies between the black and the red particles. This is the most difficult and subtle part of the argument.

4.1. *Controlling red neighbours*: *An overview.* It turns out that controlling the number of red neighbours is the easy part. Observe that the dynamics performed by the red particles during a single constant-colour phase of the chameleon process is simply a symmetric exclusion process. Thus by NA if given R_{ρ_i} (recall that ρ_i is the beginning of the ith round) the expected number of red particles neighbouring vertex v at time $\rho_i + t_{\mathrm{round}} - 1$ is at most $(\frac{r}{n} + c)d$, the (conditional) probability (given R_{ρ_i}) of having more than $(\frac{r}{n} + 2c)d$ red particles around vertex v at time $\rho_i + t_{\mathrm{round}} - 1$ can be made arbitrary small, provided d is large enough (as explained above, we may assume the degree is arbitrarily large; where $c > 0$ is some small absolute constant). Crucially, by CNA the same holds even when we condition on v being occupied by a red particle at the end of the constant-colour phase (i.e., at time $\rho_i + t_{\mathrm{round}} - 1$). This motivates considering the following set for round i:

$$\begin{aligned}&\mathrm{Nice}(i)\\&:= \left\{ v : \text{expected no. red neighbrs of } v \text{ in } t_{\mathrm{round}} - 1 \text{ time units} \leq d\left(\frac{|\mathrm{R}_{\rho_i}|}{n} + c\right)\right\},\end{aligned} \tag{41}$$

where the expectation inside the event above is conditional on R_{ρ_i}.

It suffices to control the expected number of red particles which lie in $\mathrm{Nice}(i)$ at the end of the constant-colour phase of the ith round, as by the above reasoning it is very unlikely for each such red particle to have more than $d(\frac{|\mathrm{R}_{\rho_i}|}{n} + 2c)$ red neighbours at that time. Using NA one can argue that if the last expectation is large, then the actual number of such red particles is unlikely to deviate from it by a lot. However, it turns out to not be necessary for our purposes.

To control the aforementioned (conditional) expectation (given R_{ρ_i}), we observe that the last expectation equals

$$|\mathrm{R}_{\rho_i}|\mathrm{P}_{\mathrm{Unif}(\mathrm{R}_{\rho_i})}\big[X_{t_{\mathrm{round}}-1} \in \mathrm{Nice}(i)\big]. \tag{42}$$

By Proposition 2.8 and some algebra (see Lemma 5.5 for the actual details), we deduce that if $\mathrm{P}_{\mathrm{Unif}(\mathrm{R}_{\rho_i})}[X_{t_{\mathrm{round}}-1} \in \mathrm{Nice}(i)]$ is smaller than $\pi(\mathrm{Nice}(i)) - c$, then we must have that the L_2 distance of $\mathrm{P}_{\mathrm{Unif}(\mathrm{R}_{\rho_i})}[X_{t_{\mathrm{round}}-1} \in \bullet]$ from π is proportional to $\frac{1}{\sqrt{\pi(V\setminus\mathrm{Nice}(i))}}$. By a simple counting argument (see Lemma 5.3), we must have that

$$\big|V \setminus \mathrm{Nice}(i)\big| \lesssim |\mathrm{R}_{\rho_i}|, \tag{43}$$

which means that the last L_2 distance is $\gtrsim \frac{1}{\sqrt{\pi(\mathrm{R}_{\rho_i})}} \asymp \|\mathrm{Unif}(\mathrm{R}_{\rho_i}) - \pi\|_{2,\pi}$.

In simple words, if the duration of the constant-colour relaxation phase is such that the L_2 distance from the uniform distribution of a random red particle, chosen uniformly at random, drops by the end of the phase by some sufficiently large constant factor, compared to its value at the beginning of the round (which is $\|\mathrm{Unif}(\mathrm{R}_{\rho_i}) - \pi\|_{2,\pi}$), then with a large probability (in some quantitative manner) a certain fraction of the red particles will have few red neighbours at the end of the relaxation phase (Lemma 5.5). Using the Poincaré inequality (24), it follows from our choices of the durations of the rounds that the aforementioned L_2 distance indeed drops by a constant factor, which can be made arbitrarily large by adjusting the constant C_{round}.

For the sake of being precise, we note that the above argument breaks down when $|\mathrm{R}_{\rho_i}| \wedge |\mathrm{W}_{\rho_i}| \geq \varrho_0 n$ for a certain ϱ_0 depending on the choice of c. Fortunately, in this regime we can work directly with the white particles and argue that at the end of the constant colour phase the expected number of red particles with at least $c\varrho_0 d$ white neighbours is of order n. This will be obtained as a relatively simple consequence of the Poincaré inequality, and only requires the duration of a round to be $\Omega(t_{\mathrm{rel}})$.

4.2. *Controlling black neighbours*: *An overview.* Controlling the number of black neighbours turns out to be a much harder task. By abuse of notation (treating B_t and B_{t+s} as sets), consider

$$Z_v(t,s) := \sum_u \mathbf{1}_{\{u \in \mathrm{B}_t\}} P_s\big(u, N(v)\big) = \mathbb{E}\big[\big|\mathrm{B}_{t+s} \cap N(v)\big| \big| \mathrm{B}_t\big], \tag{44}$$

where $N(x)$ is the neighbour set of vertex x. Using the NA property it is not hard to show (see Lemma 5.11) that if $(\mathbf{1}_{\{u \in \mathrm{B}_0\}} : u \in V)$ has marginals close to k/n (i.e., after a burn-in period) then $\mathrm{P}[Z_v(t,s) > (\frac{k}{n} + c)d]$ decays exponentially in $\frac{1}{\max_{x,y} P_s(x,y)}$ for all s and $v \in V$. This estimate, which is one of the key ideas in this work, is inspired from the proof of the main result in [3] (and a variant of that result whose proof also utilized NA). If $s \geq t_*(\epsilon)$, it is immediate from the definition of $t_*(\epsilon)$, that $\max_{x,y} P_s(x,y) \leq \frac{\epsilon}{\log n}$ and so this probability is $\ll n^{-20}$ for suitably chosen ϵ.

Unfortunately, this does not yield the desired conclusion, since conditioned on having a red particle at v at time $t+s$ changes the distribution of the number of black neighbours of v at that time. To overcome this difficulty, we have to take the duration of the round to be $t_{\mathrm{round}} := C_{\mathrm{round}}(t_*(\epsilon) + s_*(\epsilon) + t_{\mathrm{rel}}) + 1$, and consider two cases. We show that for each red particle, the expected number of neighbouring particles it has at the end of the constant-colour relaxation phase, which interacted with it during the first $t_*(\epsilon)$ time units of the round can be made at most cd, provided we take C_{round} to be large enough (see Lemma 5.9). This is obtained by exploiting the definition of $s_*(\epsilon)$, along with a delicate use of negative correlation. Lastly, we show that a variant of the aforementioned large deviation estimate applies to the black particles that did not interact during the first $t_*(\epsilon)$ time units of the round with the considered red particle, and that for such black particles we need not worry about the dependencies with this red particle.

5. Results to control neighbours of red particles.

5.1. *The red neighbours.* Recall that P_t is the heat-kernel of a single walk on G. We write T for $t_{\mathrm{round}} - 1$, that is, T denotes the length of a constant-colour phase. Motivated by (41) and the following paragraph we make the following definition.

DEFINITION 5.1. For each subset $S \subseteq V$, let $e_T(v,S) := \sum_{u:v\sim u} P_T(u,S)$ and define $\mathrm{Nice}(S)$ as

$$\mathrm{Nice}(S) := \left\{ v \in V : e_T(v,S) < d\left(\frac{1}{32} + \frac{|S|}{n}\right)\right\}.$$

REMARK 5.2. We will later (see Section 7.2) modify this definition by replacing adjacency with proximity. This will allow us to deal with the case of small degree.

From this definition, we see that the set $\mathrm{Nice}(S)$ consists of vertices which have "few" neighbours (in expectation) at time T which came from (at time 0) the set S. The reader should think of S as the set occupied by the red particles at the beginning of a round. In Section 6, we make use of this definition with S being the set of red vertices. Motivated by (43), we now lower-bound the size of $\mathrm{Nice}(S)$ by a simple counting argument, involving only its definition.

LEMMA 5.3. *For each $S \subseteq V$,*

$$\left|\mathrm{Nice}(S)^{\complement}\right| \le \left(\frac{1}{32} + \frac{|S|}{n}\right)^{-1} |S|.$$

PROOF. The definition of $\mathrm{Nice}(S)$ yields that $d(\frac{1}{32} + \frac{|S|}{n})|\mathrm{Nice}(S)^{\complement}| = \sum_{v \in \mathrm{Nice}(S)^{\complement}} d(\frac{1}{32} + \frac{|S|}{n})$ is

$$\le \sum_{v \in \mathrm{Nice}(S)^{\complement}} \sum_{u : v \sim u} P_T(u, S) \le \sum_{u} \sum_{v : v \sim u} P_T(u, S) = d|S|,$$

which proves the result. $\square$

REMARK 5.4. The analogous result which is used for the small degree case is Lemma 7.2.

The next lemma (motivated by (42)) gives a bound on the probability that a random walk started uniformly from set S is in $\mathrm{Nice}(S)$ at time T. The proof uses Proposition 2.8 combined with the Poincaré inequality (24).

LEMMA 5.5 (Proof in the Supplementary Material [17], Appendix C.4). *Denote the uniform distribution on S by π_S. For each $\varepsilon \in (0, 1)$, there exist $C_{5.5}(\varepsilon) > 1$ such that for all $C_{\mathrm{round}} > C_{5.5}(\varepsilon)$ and all $S \subset V$ with $2|S| \le n$,*

$$\mathrm{P}_{\pi_S}\left[X_T \in \mathrm{Nice}(S)\right] \ge \pi\left(\mathrm{Nice}(S)\right) - \varepsilon.$$

REMARK 5.6. See Lemma 9.5 for the version of this result to be used in the proof of Theorem 1.2 (mixing for sublinear number of particles).

For $S \subseteq V$, we define $N(S) := \mathrm{Nice}(S) \cap I_{[0,T]}(S)$, which are the $\mathrm{Nice}(S)$ vertices occupied at time T by particles initially in S, and further for $\theta \in (0, 1)$, we define a subset of $N(S)$ as

$$\mathrm{BN}(S)_\theta := \left\{v \in N(S) : \sum_{u : v \sim u} \mathbf{1}_{\{I_{[0,T]}^{-1}(u) \in S\}} > \theta d\right\},$$

which are the $N(S)$ vertices which have "many" ($> \theta d$) neighbours also occupied at time T by particles initially in S. Similarly, we define a set $\mathrm{GN}(S)_\theta$ to be $N(S) \setminus \mathrm{BN}(S)_\theta$ (here the B in $\mathrm{BN}(S)_\theta$ stands for "bad" and the G in $\mathrm{GN}(S)_\theta$ for "good"). We control the number of such vertices with the following lemma (think of θ below as being in $(\frac{|S|}{n} + \frac{1}{32}, \frac{|S|}{n} + \frac{1}{16}]$, and observe that for such θ we may pick $\lambda > 0$ sufficiently small such that $-\lambda\theta + (e^\lambda - 1)(\frac{1}{32} + \frac{|S|}{n}) \le -c\lambda$).

LEMMA 5.7. *For each $S \subseteq V$, $\theta \in (0, 1)$, $\lambda > 0$ and $v \in V$,*

$$\mathrm{P}\left[v \in \mathrm{BN}(S)_\theta \mid v \in N(S)\right] < \exp\left\{d\left(-\lambda\theta + (e^\lambda - 1)\left(\frac{1}{32} + \frac{|S|}{n}\right)\right)\right\}.$$

REMARK 5.8. Proving this lemma relies crucially on the CNA property. See Lemma 7.4 for the version of this result for small degree graphs.

PROOF. For each $v \in \mathrm{Nice}(S)$ and $\lambda > 0$,

$$\begin{aligned}
\mathrm{P}\big[v \in BN(S)_\theta \,|\, v \in N(S)\big] &= \mathrm{P}\Big[\sum_{u:v\sim u} \mathbf{1}_{\{I_{[0,T]}^{-1}(u)\in S\}} > \theta d \,|\, v \in I_{[0,T]}(S)\Big] \\
\text{(Chernoff)} \quad &\le e^{-\lambda\theta d}\mathbb{E}\Big[\exp\Big\{\lambda \sum_{u:v\sim u} \mathbf{1}_{\{I_{[0,T]}^{-1}(u)\in S\}}\Big\} \,|\, v \in I_{[0,T]}(S)\Big] \\
\text{(CNA then NA)} \quad &\le e^{-\lambda\theta d} \prod_{u:v\sim u} \mathbb{E}\big[\exp\{\lambda \mathbf{1}_{\{I_{[0,T]}^{-1}(u)\in S\}}\}\big] \\
&= e^{-\lambda\theta d} \prod_{u:v\sim u} \big(1 + (e^\lambda - 1)\mathrm{P}[u \in I_{[0,T]}(S)]\big) \\
(1 + x \le e^x) \quad &\le e^{-\lambda\theta d} \exp\Big\{\sum_{u:v\sim u} (e^\lambda - 1)\mathrm{P}[u \in I_{[0,T]}(S)]\Big\} \\
&= \exp\Big\{-\lambda\theta\hat{d} + (e^\lambda - 1)\sum_{u:v\sim u} P_T(u, S)\Big\} \\
(v \in \mathrm{Nice}(S)) \quad &< \exp\Big\{d\Big(-\lambda\theta + (e^\lambda - 1)\Big(\frac{1}{32} + \frac{|S|}{n}\Big)\Big)\Big\},
\end{aligned}$$

as required. $\square$

5.2. *The black neighbours.* Recall the modified graphical construction from Section 3.1. Recall also that an *interaction* occurs between two particles occupying vertices u, v in the exclusion/interchange process (constructed using the modified graphical construction as described in Section 3.1) when edge $\{u, v\}$ rings. For $a, b \in V$, and $t \ge 0$, let $N_t(a, b)$ denote the number of interactions during time interval $[0, t]$ of the particles at vertices a and b at time 0.

For each $v \in V$ and $0 \le t < T$, we also define a random variable $\hat{N}_t(v)$ to be the number of interactions during time interval $[0, t]$ of the particle at vertex v at time 0 with its time-T neighbours, that is,

$$\hat{N}_t(v) := \sum_{u:I_{[0,T]}(v)\sim u} N_t\big(v, I_{[0,T]}^{-1}(u)\big).$$

The next lemma gives control on the expected value of $\hat{N}_{t_*(\epsilon)}(v)$. We will apply this to control the expected number of black particles which interact with red particles during time interval $[0, t_*]$ for any initial configuration of black and red particles.

LEMMA 5.9. *For all $\epsilon \in (0, 1)$ and $C_{\mathrm{round}} \ge 1$, we have*

$$\max_{v\in V} \mathbb{E}\big[\hat{N}_{t_*(\epsilon)}(v)\big] \le 8d\epsilon.$$

REMARK 5.10. The proof of this lemma makes use of the NA property. See Lemma 7.5 for the version of this result for small degree graphs.

PROOF. We first write

$$\begin{aligned}
\hat{N}_{t_*(\epsilon)}(v) &= \sum_u \mathbf{1}_{\{I_{[0,T]}(v)\sim u\}} N_{t_*(\epsilon)}\big(v, I_{[0,T]}^{-1}(u)\big) \\
&= \sum_w \mathbf{1}_{\{I_{[0,T]}(v)\sim I_{[0,T]}(w)\}} N_{t_*(\epsilon)}(v, w).
\end{aligned}$$

Let $\tilde{N}_t(v,w)$ denote the amount of time particles from v and w spend adjacent during the time interval $[0,t]$. We claim that for each $w \in V$, and $0 \le t < T$,

$$\mathbb{E}[\mathbf{1}_{\{I_{[0,T]}(v)\sim I_{[0,T]}(w)\}} N_t(v,w)] = \frac{2}{d}\mathbb{E}[\mathbf{1}_{\{I_{[0,T]}(v)\sim I_{[0,T]}(w)\}} \tilde{N}_t(v,w)].$$

To see this, notice that conditionally on the unordered pair of trajectories $\{I_{[0,t]}(v), I_{[0,t]}(w)\}$, the number of times particles started from vertices v and w interact is Poisson with parameter $\frac{2}{d}\tilde{N}_t(v,w)$ (as these interactions do not affect the unordered pair of trajectories). Therefore, we have

$$\begin{aligned}
&\frac{d}{2}\mathbb{E}[\hat{N}_{t_*(\epsilon)}(v)] \\
&\quad = \sum_w \mathbb{E}[\mathbf{1}_{\{I_{[0,T]}(v)\sim I_{[0,T]}(w)\}} \tilde{N}_{t_*(\epsilon)}(v,w)] \\
&\quad = \int_0^{t_*(\epsilon)} \sum_w \mathbb{E}[\mathbf{1}_{\{I_{[0,T]}(v)\sim I_{[0,T]}(w)\}} \mathbf{1}_{\{I_{[0,s]}(w)\sim I_{[0,s]}(v)\}}]\,ds \\
&\quad = \int_0^{t_*(\epsilon)} \sum_w \sum_{a,b:a\sim b} \mathbb{E}[\mathbf{1}_{\{I_{[s,T]}(a)\sim I_{[s,T]}(b)\}} \mathbf{1}_{\{I_{[0,s]}(w)=b, I_{[0,s]}(v)=a\}}]\,ds \\
&\quad = \int_0^{t_*(\epsilon)} \sum_{a,b:a\sim b} \mathbb{E}[\mathbf{1}_{\{I_{[0,s]}(v)=a\}} \mathbf{1}_{\{I_{[s,T]}(a)\sim I_{[s,T]}(b)\}}]\,ds \\
&\quad = \int_0^{t_*(\epsilon)} \sum_{a,b:a\sim b} \mathrm{P}[I_{[0,s]}(v)=a]\mathrm{P}[I_{[s,T]}(a)\sim I_{[s,T]}(b)]\,ds \\
&\quad = \int_0^{t_*(\epsilon)} \sum_{a,b:a\sim b} \mathrm{P}[I_{[0,s]}(v)=a] \sum_{c,d:c\sim d} \mathrm{P}[I_{[s,T]}(a)=c, I_{[s,T]}(b)=d]\,ds \\
&\quad \le \int_0^{t_*(\epsilon)} \sum_{a,b:a\sim b} \mathrm{P}[I_{[0,s]}(v)=a] \\
&\qquad \cdot \sum_{c,d:c\sim d} \mathrm{P}[I_{[s,T]}(a)\in\{c,d\}]\mathrm{P}[I_{[s,T]}(b)\in\{c,d\}]\,ds,
\end{aligned}$$

where the last line follows from the NA property. Now, since $T \ge t_*(\epsilon) + s_*(\epsilon)$, for each $0 \le s \le t_*(\epsilon)$ we have that $T - s \ge s_*(\epsilon)$ and so

$$\mathrm{P}[I_{[s,T]}(b)\in\{c,d\}] \le \max_{b,c,d} \mathrm{P}[I_{[0,s_*(\epsilon)]}(b)\in\{c,d\}] \le \frac{2\epsilon}{t_*(\epsilon)}.$$

We thus obtain

$$\begin{aligned}
\mathbb{E}[\hat{N}_{t_*(\epsilon)}(v)] &\le \frac{4\epsilon}{dt_*(\epsilon)} \int_0^{t_*(\epsilon)} \sum_{a,b:a\sim b} \mathrm{P}[I_{[0,s]}(v)=a] \sum_{c,d:c\sim d} \mathrm{P}[I_{[s,T]}(a)\in\{c,d\}]\,ds \\
&\le \frac{8d\epsilon}{dt_*(\varepsilon)} \int_0^{t_*(\varepsilon)} \sum_{a,b:a\sim b} \mathrm{P}[I_{[0,s]}(v)=a]\,ds \le 8d\epsilon.
\end{aligned}$$

□

Motivated by the discussion in Section 4, for each $a,u,x,v \in V$ and $\epsilon \ge 0$, we define

$$Q(a) = Q(a,u,x,v,\epsilon) := \mathrm{P}[I_{[0,T]}(a)=u, N_{t_*(\epsilon)}(a,x)=0 \mid I_{[0,T]}(x)=v]. \tag{45}$$

The next lemma gives the large-deviation bound (for any initial configuration of black and red particles) on the number of black particles which are time-T neighbours with a red particle and which do not interact with that red particle during time interval $[0, t_*(\epsilon)]$. The proof is similar to the proof of Lemma 5.7 in that it revolves around a Chernoff bound and the NA property.

LEMMA 5.11 (Proof in the Supplementary Material [17], Appendix C.5). *Fix $\epsilon \in (0, 10^{-4}]$ and let $Q(a) = Q(a, u, x, v, \epsilon)$ be as in (45). There exists n_0 such that for all $n \geq n_0$ we have for all $2 \leq k \leq n/2$, all $u, x, v \in V$, and all $B \in (V)_{k-1}$,*

$$\sup_{s \geq t_{\mathrm{mix}}^{(\infty)}(n^{-10})} \mathrm{P}\left[\sum_{a \in V} \mathbf{1}_{\{a \in \mathrm{B}_s\}} Q(a) > \frac{k}{n} + \frac{1}{16} | \mathrm{B}_0 = B\right] \leq n^{-13}.$$

REMARK 5.12. The large deviation bound on the black particle measure needed for the proofs of Theorems 1.2 (mixing for sublinear number of particles) and 1.3 (mixing for graphs of high degree) is Lemma 7.6.

6. Loss of red in a round: Proof of Proposition 3.3 for $d \geq 10^4$. In this section, we prove Proposition 3.3 for $d \geq 10^4$. We begin with some new definitions. For each $a \in V$, let ϕ_a be the first time of the form $\tau_j \in (T, T+1)$ for which $a \in e_j$ (setting $\phi_a = \infty$ if no such time exists). If $\phi_a < \infty$, then define $F_a = I^{-1}_{(T, \phi_a)}(b)$ where b is the other vertex on edge e_j; if instead $\phi_a = \infty$ then we write $F_a = *$. (This notation is similar to that appearing in [39], Section 9.2.) Recall also the definition of an (α, t)-good configuration from Definition 3.1. We determine the kinds of configurations that are (α, T)-good.

LEMMA 6.1. *Suppose $d \geq 10^4$. If $C_{\mathrm{round}} > C_{5.5}(10^{-4})$ then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process satisfying*

$$\max_{b,v,a} \sum_{z \in \mathrm{B}} Q(z, b, v, a, 10^{-4}) \leq \frac{k}{n} + \frac{1}{16},$$

is (α_1, T)-good, for $T = C_{\mathrm{round}}(t_{\mathrm{rel}} + t_*(10^{-4}) + s_*(10^{-4}))$ and some $\alpha_1 > 0$.

PROOF. Recall the definition of H_t from §3.1. Without loss of generality, suppose $|R| \leq |W|$. We bound H_T by only counting pink particles created from red and white particles satisfying: the red particle is on some vertex a at time T and the white on some vertex b with $a \sim b$, and $\phi_a = \phi_b < \infty$. Observe that we have

$$\begin{aligned} H_T &\geq \sum_{b \in I_{[0,T]}(W)} \mathbf{1}_{\{\bigcup_{a \in I_{[0,T]}(R)} \{F_a = b, \phi_a = \phi_b\}\}} \\ &= \sum_{b \in I_{[0,T]}(W)} \sum_{a \in I_{[0,T]}(R)} \mathbf{1}_{\{F_a = b, \phi_a = \phi_b\}}, \end{aligned}$$

where the equality follows from the fact that the events $\{F_a = b, \phi_a = \phi_b\}$ are disjoint. Recall the definitions of $N(R)$ (as a subset of the Nice(R) vertices) and GN(R) (as the subset of $N(R)$ which are "good") from the discussion after Lemma 5.5. Taking an expectation in the above inequality gives, for any $\theta \in (0, 1)$ and $M = (B, R, W)$,

$$\begin{aligned} \mathbb{E}_M[H_T] &\geq \sum_{a,b: a \sim b} \mathrm{P}[a \in I_{[0,T]}(R), b \in I_{[0,T]}(W), F_a = b, \phi_a = \phi_b] \\ &= \sum_{a,b: a \sim b} \mathrm{P}[a \in I_{[0,T]}(R), b \in I_{[0,T]}(W)] \mathrm{P}[F_a = b, \phi_a = \phi_b] \qquad (46) \\ &\geq \sum_{a,b: a \sim b} \mathrm{P}[a \in GN(R)_\theta, b \in I_{[0,T]}(W)] \mathrm{P}[F_a = b, \phi_a = \phi_b], \end{aligned}$$

where the second equality follows by independence of the edge-rings before and after time T. Notice now that we have

$$\begin{aligned}\mathrm{P}[F_a = b, \phi_a = \phi_b] &= \mathrm{P}[F_a = b, \phi_a = \phi_b | F_a \neq *]\mathrm{P}[F_a \neq *] \\ &= \frac{1}{2d-1}\mathrm{P}[F_a \neq *] \geq \frac{1}{4d},\end{aligned}$$

where the inequality follows from the fact that some edge incident to vertex a will ring during time interval $(T, T+1)$ with probability $1 - e^{-1} > 1/2$. Plugging this into (46) gives

$$\mathbb{E}_M[H_T] \geq \frac{1}{4d} \sum_{a,b:a\sim b} \mathrm{P}\big[a \in GN(R)_\theta, b \in I_{[0,T]}(W)\big]. \tag{47}$$

Instead of considering pairs of red and white particles, we consider pairs of red and red, and pairs of red and black. So we now decompose

$$\begin{aligned}&\mathrm{P}\big[a \in GN(R)_\theta, b \in I_{[0,T]}(W)\big] \\ &\quad = \mathrm{P}\big[a \in GN(R)_\theta\big]\big(1 - \mathrm{P}\big[b \in I_{[0,T]}(R) \mid a \in GN(R)_\theta\big]\big) \\ &\qquad - \mathrm{P}\big[a \in GN(R)_\theta, b \in I_{[0,T]}(B)\big].\end{aligned} \tag{48}$$

Using Lemma 5.7 we have, for any $\theta \in (0,1)$ and $\lambda > 0$, the bound

$$\mathrm{P}\big[a \in GN(R)_\theta\big] \geq \big(1 - L\big(\lambda, \theta, d, |R|\big)\big)\mathrm{P}\big[a \in N(R)\big], \tag{49}$$

where $L(\lambda, \theta, d, r) := \exp\{-\lambda\theta d + (e^\lambda - 1)(\frac{1}{32} + \frac{r}{n})d\}$.

We decompose the term $\mathrm{P}[a \in \mathrm{GN}(R)_\theta, b \in I_{[0,T]}(B)]$ according to the starting location of particle at vertex a at time T:

$$\begin{aligned}&\mathrm{P}\big[a \in GN(R)_\theta, b \in I_{[0,T]}(B))\big] \leq \mathrm{P}\big[a \in N(R), b \in I_{[0,T]}(B)\big] \\ &\quad = \sum_{v\in R} \mathrm{P}\big[a \in \mathrm{Nice}(R), b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big] \\ &\quad = \sum_{v\in R} \mathbf{1}_{\{a\in\mathrm{Nice}(R)\}}\mathrm{P}\big[b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big]\end{aligned} \tag{50}$$

where in the last line we have used the fact that being in Nice is a deterministic property.

Using the definition of $\mathrm{GN}(R)_\theta$ for any $\theta \in (0,1)$ and $\lambda > 0$ and combining equations (47)–(50), we obtain

$$\begin{aligned}\mathbb{E}_M[H_T] \geq{}& \frac{1}{4d}\sum_a \mathrm{P}\big[a \in N(R)\big]\big(1 - L\big(\lambda, \theta, d, |R|\big)\big)(d - \theta d) \\ &- \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in\mathrm{Nice}(R)\}}\mathrm{P}\big[b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big].\end{aligned} \tag{51}$$

We now further decompose $\mathrm{P}[b \in I_{[0,T]}(B), a = I_{[0,T]}(v)]$ into two terms, depending on whether the trajectories of particles started from vertices a and b are adjacent, and use Markov's inequality to give

$$\begin{aligned}&\mathrm{P}\big[b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big] \\ &\quad \leq \mathrm{P}\big[N_{t_*(10^{-4})}\big(I^{-1}_{[0,T]}(a), I^{-1}_{[0,T]}(b)\big) = 0, b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big] \\ &\qquad + \mathbb{E}\big[N_{t_*(10^{-4})}\big(I^{-1}_{[0,T]}(a), I^{-1}_{[0,T]}(b)\big)\mathbf{1}_{\{a=I_{[0,T]}(v)\}}\big].\end{aligned} \tag{52}$$

Combining equations (51) and (52), we obtain for any $\theta \in (0,1)$ and $\lambda > 0$,

$$
\begin{aligned}
&\mathbb{E}_M[H_T] \\
&\quad \geq \frac{1}{4}\sum_a \mathrm{P}\big[a \in N(R)\big]\big(1 - L\big(\lambda, \theta, d, |R|\big)\big)(1-\theta) \\
&\qquad - \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \mathrm{P}\big[N_{t_*(10^{-4})}\big(I_{[0,T]}^{-1}(a), I_{[0,T]}^{-1}(b)\big) = 0, \\
&\qquad b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big] \\
&\qquad - \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \mathbb{E}\big[N_{t_*(10^{-4})}\big(I_{[0,T]}^{-1}(a), I_{[0,T]}^{-1}(b)\big)\mathbf{1}_{\{a = I_{[0,T]}(v)\}}\big].
\end{aligned}
\tag{53}
$$

For the second term on the right-hand side, we have

$$
\begin{aligned}
&\frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \\
&\qquad \cdot \mathrm{P}\big[N_{t_*(10^{-4})}\big(I_{[0,T]}^{-1}(a), I_{[0,T]}^{-1}(b)\big) = 0, b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big] \\
&\quad = \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \\
&\qquad \cdot \sum_{z\in B} \mathrm{P}\big[N_{t_*(10^{-4})}(v,z) = 0, b = I_{[0,T]}(z), a = I_{[0,T]}(v)\big] \\
&\quad = \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \mathrm{P}\big[a = I_{[0,T]}(v)\big]\sum_{z\in B} Q\big(z,b,v,a,10^{-4}\big) \\
&\quad \leq \left(\frac{k}{n} + \frac{1}{16}\right)\cdot \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \mathrm{P}\big[a = I_{[0,T]}(v)\big] \\
&\quad = \left(\frac{k}{n} + \frac{1}{16}\right)\cdot \frac{1}{4}\sum_a \mathrm{P}\big[a \in N(R)\big],
\end{aligned}
\tag{54}
$$

where the inequality follows from the assumption on the configuration M.

The third term on the right-hand side of (53) is

$$
\begin{aligned}
&\frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}} \mathbb{E}\big[N_{t_*(10^{-4})}\big(I_{[0,T]}^{-1}(a), I_{[0,T]}^{-1}(b)\big)\mathbf{1}_{\{a = I_{[0,T]}(v)\}}\big] \\
&\quad \leq \frac{1}{4d}\sum_{v\in R} \mathbb{E}\Bigg[\sum_a \mathbf{1}_{\{a = I_{[0,T]}(v)\}} \sum_{b:b\sim I_{[0,T]}(v)} N_{t_*(10^{-4})}\big(v, I_{[0,T]}^{-1}(b)\big)\Bigg] \\
&\quad = \frac{1}{4d}\sum_{v\in R} \mathbb{E}\Bigg[\sum_{b:b\sim I_{[0,T]}(v)} N_{t_*(10^{-4})}\big(v, I_{[0,T]}^{-1}(b)\big)\Bigg] \\
&\quad = \frac{1}{4d}\sum_{v\in R} \mathbb{E}\big[\hat{N}_{t_*(10^{-4})}(v)\big] \leq \sum_{v\in R} 2\times 10^{-4} = 2\times 10^{-4}|R|,
\end{aligned}
\tag{55}
$$

where the second inequality follows from Lemma 5.9. Plugging equations (54) and (55) into (53) gives, for any $\theta \in (0,1)$ and $\lambda > 0$,

$$\begin{aligned}(56)\qquad &\mathbb{E}_M[H_T] \\ &\geq \frac{1}{4}\sum_a \mathrm{P}\big[a \in N(R)\big]\bigg\{\big(1 - L\big(\lambda,\theta,d,|R|\big)\big)(1-\theta) - \frac{k}{n} - \frac{1}{16}\bigg\} - 2 \times 10^{-4}|R|.\end{aligned}$$

Choosing $\lambda = 0.05$, $\theta = \frac{9}{16} - \frac{k}{2n}$ and using the bound $|R|/n \leq \frac{1}{2} - \frac{k}{2n}$, we have that $-\lambda\theta + (e^\lambda - 1)(\frac{1}{32} + \frac{|R|}{n}) \leq -\lambda(\frac{9}{16} - \frac{1}{32} - \frac{1}{2}) + (\frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \lambda^4)(\frac{1}{32} + \frac{1}{2})$, where we have used $\lambda \leq e^\lambda - 1 \leq \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \lambda^4$ for $\lambda \in [0, 0.05]$, and so

$$\frac{1}{d}\log L\big(\lambda,\theta,d,|R|\big) = -\lambda\theta + \big(e^\lambda - 1\big)\big(1/32 + |R|/n\big) \leq -0.0008,$$

and so since $d \geq 10^4$, we obtain

$$\big(1 - L\big(\lambda,\theta,d,|R|\big)\big)(1-\theta) - \frac{k}{n} - \frac{1}{16} \geq -\frac{k}{n}\bigg(1 - \frac{1-e^{-8}}{2}\bigg) + \frac{6-7e^{-8}}{16} > \frac{1}{16}.$$

Plugging this into (56) gives the bound

$$(57)\qquad \mathbb{E}_M[H_T] \geq \frac{1}{64}\mathbb{E}\big[\big|N(R)\big|\big] - 2 \times 10^{-4}|R|.$$

Notice now that $\mathbb{E}[|N(R)|] = |R|\mathrm{P}_{\pi_R}(X_T \in \mathrm{Nice}(R))$, for (X_t) a realisation of $\mathrm{RW}(G)$, and so by Lemmas 5.3 and 5.5 we have that, since $C_{\text{round}} > C_{5.5}(10^{-4})$, $\mathbb{E}[|N(R)|] \geq |R|(\frac{1}{17} - 10^{-4})$. Hence from (57) we obtain the bound

$$\mathbb{E}_M[H_T] \geq |R|\bigg(\frac{1}{1088} - \frac{129}{64} \times 10^{-4}\bigg) > 0.0007|R|.$$

The proof is completed by taking any $\alpha_1 \leq 0.0007$. $\square$

PROOF OF PROPOSITION 3.3 FOR $d \geq 10^4$. Recall the notation $t_0 = t_{\text{mix}}^{(\infty)}(n^{-10})$ and let $t \geq t_0$. By Lemma 5.11, we have that for any $B \in (V)_{k-1}$, and n sufficiently large, by a union bound

$$\begin{aligned}&\mathrm{P}\bigg[\max_{b,v,a}\sum_{z \in \mathrm{B}_t} Q\big(z,b,v,a,10^{-4}\big) \leq \frac{k}{n} + \frac{1}{16}\big|\mathrm{B}_0 = B\bigg] \\ &\quad\geq 1 - \sum_{b,v,a}\mathrm{P}\bigg[\sum_{z \in B_t} Q\big(z,b,v,a,10^{-4}\big) > \frac{k}{n} + \frac{1}{16}\big|\mathrm{B}_0 = B\bigg] \geq 1 - n^{-10}.\end{aligned}$$

Therefore, if we have $C_{\text{round}} > C_{5.5}(10^{-4})$ then, by Lemma 6.1, since $d \geq 10^4$, with probability at least $1 - n^{-10}$, M_t (the configuration of the chameleon process at time t) is (α_1, T)-good, for $T = C_{\text{round}}(t_{\text{rel}} + t_*(10^{-4}) + s_*(10^{-4}))$ and some $\alpha_1 > 0$, that is, $\beta(\alpha_1, t_{\text{round}} - 1) \leq n^{-10}$.

This completes the proof taking $\alpha = \alpha_1$. $\square$

7. Modifications to the main approach.

7.1. *Generalising the results of Section 5.* For the case of $d < 10^4$ for Theorem 1.1 (and for other values of d in general), it will be useful to artificially inflate the degree of vertices by adding "dummy" directed edges (of zero weight) to the graph (without the addition of new vertices). The number of edges we need to add varies according to the values of k and d and we let $\hat{d}$ denote the new out-degree of all of the vertices (which is the number of undirected edges plus the number of directed out-edges from a vertex). We will always add these edges between vertices within graph distance at most $\hat{d}$ in the original graph. These edges are assigned weight 0 and so never ring and play no role in the dynamics of the processes, instead just affecting the structure of the graph (in particular adjacency).

Any such graph that has these additional edges is referred to as a *modified* graph and we write $v \overset{\to}{\sim} u$ to indicate that either (v, u) or $\{v, u\}$ is an edge in a modified graph. We denote the maximal in-degree in the modified graph by $d_{\max}^{\text{in}}$.

If we could ensure that the in-degrees were all equal (to the out-degree), the argument from Sections 5–6 would work almost verbatim (the main difference is that one has to replace adjacency by adjacency in the modified graph, apart from when controlling the number of interactions between a pair of particles). If the graph is vertex-transitive one can easily ensure this. Alas, in general one cannot do this. As we inflate the degree only when $d < 10^4$, it follows that there exists an absolute constant D such that $d_{\max}^{\text{in}} \le D\hat{d}$. We now explain how, from a high-level perspective, this leads only to minor changes in the outline of the argument. While we expect most readers to be satisfied with this outline, we present all details of the proof below.

Recall that when $d \ge 10^4$ the (indirect) argument for controlling the expected number of white neighbours that red particles have by considering the number of red and black neighbours they have breaks down when $|R| \wedge |W| > \varrho_0 n$ for a certain constant ϱ_0. Fortunately, in this regime we could work directly with the white particles. In the case $d < 10^4$, the indirect argument breaks down for a smaller value of ϱ_0, due to an additional factor $D \ge \frac{d_{\max}^{\text{in}}}{\hat{d}}$ appearing in analogous statements to ones from Sections 5–6 related to the analysis of the number of red neighbours. This is not a problem, as when $|R| \wedge |W| > \varrho_0 n$ (and $d < 10^4$) we argue that at the end of the constant colour phase, the expected number of red particles with at least one white neighbour in the modified graph is of order n (provided $T \gtrsim t_{\text{rel}}$, where $T + 1$ is the duration of the round). As the modified graph has bounded degree, this suffices to argue that $\mathbb{E}[H_T] \gtrsim n$, as required.

Recall the definition of Nice from Section 5.1. We modify this definition to deal with these modified graphs and in the sequel this is the definition of Nice that we use (i.e., every future use of Nice refers to this new definition).

DEFINITION 7.1. For each subset $S \subseteq V$, let $e_T(v, S) := \sum_{u : v \overset{\to}{\sim} u} P_T(u, S)$ and define $\text{Nice}(S)$ as

$$\text{Nice}(S) := \left\{ v \in V : e_T(v, S) < \hat{d}\left(\frac{1}{32} + \frac{|S|}{n}\right)\right\}.$$

The equivalent statement of Lemma 5.3 is the following. We omit the proof as it follows in a similar manner (i.e., a simple counting argument together with the fact that the modified graph has out-degree $\hat{d}$ at each site).

LEMMA 7.2 (Proof omitted). *For each $S \subseteq V$,*

$$\left|\text{Nice}(S)^{\complement}\right| \le \left(\frac{1}{32} + \frac{|S|}{n}\right)^{-1} \frac{d_{\max}^{\text{in}}}{\hat{d}} |S|.$$

REMARK 7.3. Using Lemma 7.2, it is possible to prove (proof omitted) that Lemma 5.5 still holds with the new definition of Nice and so will make use of this lemma in this section.

Next, we recall the definition of $\mathrm{BN}(S)$ for $S \subseteq V$ from §5.1 and redefine it for modified graphs as

$$\mathrm{BN}(S)_\theta := \Big\{ v \in N(S) : \sum_{u : v \overset{\to}{\sim} u} \mathbf{1}_{\{I_{[0,T]}^{-1}(u) \in S\}} > \theta \hat{d} \Big\}.$$

We also redefine set $\mathrm{GN}(S)_\theta$ to be $N(S) \setminus \mathrm{BN}(S)_\theta$. The analogue of Lemma 5.7 is the following.

LEMMA 7.4. *For each $S \subseteq V$, $\theta \in (0, 1)$, $\lambda > 0$ and $v \in V$,*

$$\mathrm{P}\big[v \in \mathrm{BN}(S)_\theta \mid v \in N(S)\big] < \exp\bigg\{ \hat{d} \bigg(-\lambda\theta + (e^\lambda - 1) \bigg(\frac{1}{32} + \frac{|S|}{n} \bigg) \bigg) \bigg\}.$$

Next, we recall Lemma 5.9 which bounds the expected value of $\hat{N}_t(v)$. We redefine this quantity in terms of modified graphs as follows:

$$\hat{N}_t(v) := \sum_{u : I_{[0,T]}(v) \overset{\to}{\sim} u} N_t\big(v, I_{[0,T]}^{-1}(u)\big).$$

We present below the version of Lemma 5.9 to be used for modified graphs. The proof is omitted and we remark that the main difference in the proof of this result compared with Lemma 5.9 is that we define $\tilde{N}_t(v, w)$ to be the amount of time particles from v and w spend adjacent w.r.t. G (as opposed to w.r.t. the modified graph) during the time interval $[0, t]$.

LEMMA 7.5 (Proof omitted). *For all $\epsilon \in (0, 1)$ and $C_{\mathrm{round}} \geq 1$, we have*

$$\max_{v \in V} \mathbb{E}\big[\hat{N}_{t_*(\epsilon)}(v)\big] \leq 4\epsilon\big(d_{\max}^{\mathrm{in}} + \hat{d}\big).$$

We now show that after a burn-in period we have a large deviation estimate of the black particle measure. After a burn-in period, the occupation by the black particle measure has marginals extremely close to k/n and has the NA property. A simple calculation involving the Laplace transform (Lemma 7.6) shows that it satisfies large deviation estimates similar to the ones available in the independent case. From this, along with a union bound, one can derive (Corollary 7.7) that at each given time after a burn in period, the probability of having a configuration satisfying that given this current configuration, the probability of having more than $(\frac{k}{n} + c)d$ black neighbours of a vertex after T additional time units (where $T + 1$ is the duration of a round) is $\ll n^{-10}$ (i.e., if we start a round at this time, the probability that at the end of the constant-colour phase we have at least $(\frac{k}{n} + c)d$ black neighbours is small).

The proof is similar to the proof of Lemma 5.7. For $\varepsilon \in (0, 1)$, $n \in \mathbb{N}$ and $2 \leq k \leq n/2$, we denote $m_{\varepsilon,n,k} := \max\{\log \frac{\varepsilon n}{e^2 k}, \frac{\varepsilon n}{2k}(\frac{1}{2} - \frac{\varepsilon n}{k})\}$.

LEMMA 7.6 (Proof in the Supplementary Material [17], Appendix C.6). *Fix $\varepsilon \in (0, 1)$. There exists $n_0 = n_0(\varepsilon)$ such that for all $n \geq n_0$, $2 \leq k \leq n/2$, $B \in (V)_{k-1}$, $v \in V$, and $s \geq t_{\mathrm{mix}}^{(\infty)}(n^{-10})$,*

$$\mathrm{P}\bigg[\sum_{u : v \overset{\to}{\sim} u} \mathbf{1}_{\{u \in \mathrm{B}_s\}} \geq \bigg(\frac{k}{n} + \varepsilon\bigg) \hat{d} \,|\, \mathrm{B}_0 = B \bigg] \leq \exp(-\hat{d}\varepsilon m_{\varepsilon,n,k}).$$

COROLLARY 7.7. *Fix* $\varepsilon \in (0,1)$ *and for each* $t > 0$ *let* $\mathcal{F}_t$ *denote the* σ*-algebra generated by* B_t. *There exists* $n_0 = n_0(\varepsilon)$ *such that for all* $n \geq n_0$, $2 \leq k \leq n/2$, $B \in (V)_{k-1}$, $v \in V$ *and* $s_2 \geq s_1 \geq t_{\mathrm{mix}}^{(\infty)}(n^{-10})$,

$$\mathrm{P}\biggl[\mathrm{P}\biggl[\sum_{u : v \overset{\rightarrow}{\sim} u} \mathbf{1}_{\{u \in \mathrm{B}_{s_2}\}} \geq \Bigl(\frac{k}{n} + \varepsilon\Bigr)\hat{d} \,|\, \mathcal{F}_{s_1}\biggr] \geq \exp\Bigl(-\frac{1}{2}\hat{d}\varepsilon m_{\varepsilon,n,k}\Bigr)\biggr]$$
$$\leq \exp\Bigl(-\frac{1}{2}\hat{d}\varepsilon m_{\varepsilon,n,k}\Bigr).$$

PROOF. The proof immediately follows from Lemma 7.6 using Markov's inequality. $\square$

REMARK 7.8. The above corollary also holds for sufficiently small $\hat{c} \in (0,1)$ taking $s_2 \geq s_1 \geq t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$. This follows from the fact that Lemma 7.6 holds when n^{-10} is replaced with a sufficiently small $\hat{c}$ and so in particular holds when replaced with $\hat{c}/k$.

For graphs with sufficiently small degree, once the number of red particles is at least some fraction of n, it turns out that we can avoid analysing the number of red and black neighbours of a vertex (conditioned on being red), and instead directly lower bound the number of white neighbours (in fact in this case we do not even need burn-in periods). To see why, observe that the number of red particles without a nearby white particle after the relaxation phase is comparable (as $|R| \asymp n$) to the number of vertices without a nearby white particle at this time. This can be controlled with a simple argument making use of the Poincaré inequality; see Lemma 7.9. For the remaining red vertices in the proximity of a white particle, we can easily lower-bound the probability of their interaction during a unit time interval.

For a subset $S \subseteq V$, we define another subset $Q \subseteq V$ in the following way:

$$Q(S) = \biggl\{v \in V : \sum_{u : v \overset{\rightarrow}{\sim} u} P_T(u, S) < \hat{d}/16\biggr\}.$$

The reader should think of S as the set occupied by the white particles at the beginning of a round. Recall that w.l.o.g. we always consider in Sections 5–7 the case that there are as many white particles as there are red, and so $|S|/n \geq 1/4$.

We achieve control on the number of white neighbours via the following lemma.

LEMMA 7.9. *For any* $\varepsilon \in (0,1)$ *and any* $S \subset V$ *with* $|S|/n \geq 1/4$, *if* $T \geq t_{\mathrm{rel}}|\log(1/\varepsilon)|$, *then* $|Q(S)| \leq 8\varepsilon n \frac{d_{\max}^{\mathrm{in}}}{\hat{d}}$.

PROOF. If $T \geq t_{\mathrm{rel}}|\log(1/\varepsilon)|$, then since $|S|/n \geq 1/4$, the L_2-distance of $P_{\pi_S}(X_T \in \bullet)$ from π is at most 2ε by the Poincaré inequality (24), and hence this is also a bound on the L_1-distance. Therefore, by a simple counting argument and reversibility

$$\bigl|\{u : P_T(u, S) < |S|/(2n)\}\bigr| < 4\varepsilon n. \tag{58}$$

We prove the statement of the lemma by contradiction. So suppose $|Q(S)| > 8\varepsilon n \frac{d_{\max}^{\mathrm{in}}}{\hat{d}}$, that is, there are more than $8\varepsilon n \frac{d_{\max}^{\mathrm{in}}}{\hat{d}}$ vertices v for which we have $\sum_{u : v \overset{\rightarrow}{\sim} u} P_T(u, S) < \hat{d}/16 \leq \hat{d}|S|/(4n)$. Then for each $v \in Q(S)$, we must have at least $\hat{d}/2$ vertices u such that $v \overset{\rightarrow}{\sim} u$ with $P_T(u, S) < |S|/(2n)$. Now each u has in-degree at most $d_{\max}^{\mathrm{in}}$, and thus overall there are at least $\hat{d}|Q(S)|/(2d_{\max}^{\mathrm{in}})$ vertices $u \in V$ with $P_T(u, S) < |S|/(2n)$, but since we assume $|Q(S)| > 8\varepsilon n \frac{d_{\max}^{\mathrm{in}}}{\hat{d}}$, this number of vertices is at least $4\varepsilon n$. This is in contradiction with (58). $\square$

LEMMA 7.10. *Let $S \subset V$. For each $v \in Q(S)^{\complement}$,*

$$\mathrm{P}\Big[\sum_{u:v\overset{\to}{\sim}u} \mathbf{1}_{\{u\in S_T\}}=0\Big] \le \Big(\frac{31}{32}\Big)^{\hat{d}/32}.$$

REMARK 7.11. The proof of this lemma relies on the NA property.

PROOF. Notice that, since $v \in Q(S)^{\complement}$, we must have at least $\hat{d}/32$ vertices u with $v \overset{\to}{\sim} u$ such that $\mathrm{P}[u \in S_T] \ge 1/32$. Hence by the NA property,

$$\mathrm{P}\Big[\sum_{u:v\overset{\to}{\sim}u} \mathbf{1}_{\{u\in S_T\}}=0\Big] \le \prod_{u:v\overset{\to}{\sim}u} \mathrm{P}[u \notin S_T] \le \Big(\frac{31}{32}\Big)^{\hat{d}/32}.$$

□

7.2. *Proof of Proposition* 3.3 *for $d < 10^4$.* In order to prove Proposition 3.3 for $d < 10^4$, we split into two cases depending on the value of $|R| \wedge |W|$ (the minimum of the number of reds and whites in the initial configuration of the chameleon process) and a constant $\varrho_0 \in (0, 1/4)$ to be later determined. The following lemma will be used for the case $|R| \wedge |W| \ge \varrho_0 n$.

LEMMA 7.12. *Let $\varrho \in (0, 1/4)$, $C_* \ge 1$ and consider case $d < C_* \log(1/\varrho)$. There exists a constant C_*^0 such that if $C_* \ge C_*^0$ then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process with $|R| \wedge |W| \ge \varrho n$ is (α_4, T)-good for $T \ge C_{\mathrm{round}} t_{\mathrm{rel}}$ with C_{round} and α_4 depending only on ϱ and C_*.*

PROOF. We inflate the degree so that $\hat{d} = \lceil C_* \log(1/\varrho)\rceil$. Without loss of generality suppose $|R| \le |W|$. Notice that since $k \le n/2$, we have that $|W|/n \ge 1/4$.

Notice that a white particle will get pinkened during $(T, T+1)$ if there exists a red particle such that:

1. the red particle is on some vertex a at time T with a belonging to a sparse set A, and the white on some vertex b, with $a \overset{\to}{\sim} b$,
2. $\phi_a < \infty$ (i.e., vertex a is on a ringing edge during time interval $(T, T+1)$),
3. at time ϕ_a the other vertex a' incident to the ringing edge is occupied by the white particle (which may have turned pink by this time),
4. during time interval $[T, \phi_a)$ the white particle moves along a shortest trajectory from b to a'.

We remark that this will only result in pink particles being created *at time* ϕ_a if the white particle is in fact still white at time ϕ_{a-} (and otherwise it gets pinkened prior to this time). We choose the set A to have minimal size while satisfying $\sum_{a\in A} \mathrm{P}[a \in I_{[0,T]}(R)] \ge \hat{d}^{-2\hat{d}}|R|$ and with the property that no two elements of A are within graph distance (in the original graph) $2\hat{d}$. It can be shown (e.g., with a greedy construction) that $|A| \le \hat{d}^{-2\hat{d}} n$.

Observe that we can bound

$$\begin{aligned} H_T &\ge \sum_{b\in I_{[0,T]}(W)} \mathbf{1}_{\{\bigcup_{a\in I_{[0,T]}(R)\cap A}\{F_a=b,a\overset{\to}{\sim}b\}\}} \\ &= \sum_{b\in I_{[0,T]}(W)} \sum_{a\in I_{[0,T]}(R)\cap A} \mathbf{1}_{\{F_a=b,a\overset{\to}{\sim}b\}}, \end{aligned}$$

where the equality follows from the fact that each $b \in V$ is adjacent to at most one $a \in A$. Taking an expectation gives

$$\begin{aligned}\mathbb{E}_M[H_T] &\geq \sum_{a\in A}\sum_{b:a\overset{\to}{\sim}b} \mathrm{P}\big[a \in I_{[0,T]}(R), b \in I_{[0,T]}(W), F_a = b\big]\\ &= \sum_{a\in A}\sum_{b:a\overset{\to}{\sim}b} \mathrm{P}\big[a \in I_{[0,T]}(R), b \in I_{[0,T]}(W)\big]\mathrm{P}[F_a = b],\end{aligned}$$

where the equality follows by independence of the edge-rings before and after time T.

To lower bound the probability $\mathrm{P}[F_a = b]$ we fix a particular trajectory the white particle must follow, from its position at time T (vertex b) to a vertex (denoted a') adjacent to a. The trajectory chosen is one of shortest length between a and b. We additionally impose the condition that the particle must follow this trajectory during time interval $[T, T + 1/2]$. Since the degree of each vertex is less than $\hat{d}$, and vertex b is within graph distance (in the original graph) $\hat{d}$ from a, this event has probability bounded from below by some constant $c_1 > 0$ (uniformly over a and b). The event $\{F_a = b\}$ will then be satisfied if the first edge incident to vertex a to ring during $(T, T + 1)$ is edge $\{a, a'\}$ and this edge first rings during time interval $(T + 1/2, T + 1]$, an event of probability $c_2 > 0$. Hence we obtain the bound $\mathrm{P}[F_a = b] \geq c_1c_2$. Note that these constants depend on ϱ since $\hat{d}$ depends on ϱ.

Hence we have

$$\begin{aligned}\mathbb{E}_M[H_T] &\geq c_1c_2\sum_{a\in A}\sum_{b:a\overset{\to}{\sim}b} \mathrm{P}\big[a \in I_{[0,T]}(R), b \in I_{[0,T]}(W)\big]\\ &\geq c_1c_2\sum_{a\in A} \mathrm{P}\big[a \in I_{[0,T]}(R), \exists b \in I_{[0,T]}(W) : a \overset{\to}{\sim} b\big].\end{aligned}$$

Recall the definition of $Q(S)$ from Section 7.1. Decomposing the above sum (and writing $a \overset{\to}{\nsim} b$ to indicate that it is not the case that $a \overset{\to}{\sim} b$) we have

$$\begin{aligned}\mathbb{E}_M[H_T] \geq{}& c_1c_2\sum_{a\in A}\mathrm{P}\big[a \in I_{[0,T]}(R)\big]\\ &- c_1c_2\sum_{a\in Q(W)} \mathrm{P}\big[a \in I_{[0,T]}(R), \forall b \in I_{[0,T]}(W), a \overset{\to}{\nsim} b\big]\\ &- c_1c_2\sum_{a\in A\cap Q(W)^{\complement}} \mathrm{P}\big[a \in I_{[0,T]}(R), \forall b \in I_{[0,T]}(W), a \overset{\to}{\nsim} b\big]\\ \geq{}& \frac{c_1c_2}{\hat{d}^{2\hat{d}}}|R| - c_1c_2\big|Q(W)\big| - c_1c_2\sum_{a\in A\cap Q(W)^{\complement}} \mathrm{P}\big[\forall b \in I_{[0,T]}(W), a \overset{\to}{\nsim} b\big].\end{aligned}$$

By Lemma 7.9 with $\varepsilon = (\hat{d}\varrho)/(32d^{\mathrm{in}}_{\max}\hat{d}^{2\hat{d}})$, since $|W|/n \geq 1/4$, if $C_{\mathrm{round}} > \log(1/\varepsilon)$ then $|Q(W)| \leq \varrho n/(4\hat{d}^{2\hat{d}})$. Notice that for a fixed choice of $\hat{d}$, there exists a universal (over G) constant D such that $d^{\mathrm{in}}_{\max} \leq D\hat{d}$, and hence C_{round} depends only on ϱ and the choice of C_*. By Lemma 7.10, if we take $C^0_* = 2500$ then since $\varrho < 1/4$ we have that for each $a \in Q(W)^{\complement}$, $\mathrm{P}[\forall b \in I_{[0,T]}(W), a \overset{\to}{\nsim} b] \leq \varrho/4$. Hence we obtain

$$\mathbb{E}_M[H_t] \geq \frac{c_1c_2}{\hat{d}^{2\hat{d}}}|R| - \frac{c_1c_2}{\hat{d}^{2\hat{d}}}\frac{\varrho n}{4} - c_1c_2|A|\frac{\varrho}{4} \geq \frac{c_1c_2}{\hat{d}^{2\hat{d}}}\frac{|R|}{2},$$

which completes the proof with $\alpha_4 = \frac{1}{2}c_1c_2\hat{d}^{-2\hat{d}}$. $\square$

Now we consider how to deal with the case $|R| \wedge |W| < \varrho_0 n$. Recall Lemma 6.1 from Section 6. In order to prove the equivalent statement for the case $d < 10^4$, we follow a similar argument but also make use of degree-inflation. We first state a preliminary lemma which states that we can find a sparse subset of $\mathrm{Nice}(S)$ which picks-up a fraction of the time-T mass of a random walk started uniformly on S.

LEMMA 7.13 (Proof in the Supplementary Material [17], Appendix C.7). *Suppose $d < 10^4$. For any $S \subseteq V$, there exists a constant $c_{\mathrm{frac}} > 0$ and a subset $A(S)$ of $\mathrm{Nice}(S)$ such that no two members of $A(S)$ are within graph distance of 2×10^4 and such that*

$$\sum_{u \in A(S)} \mathrm{P}_{\pi_S}[X_T = u] \geq c_{\mathrm{frac}} \sum_{u \in \mathrm{Nice}(S)} \mathrm{P}_{\pi_S}[X_T = u].$$

Notice that, due to the sparseness property of $A(S)$, in the modified graph if $v, w \in A$ and $v \overset{\to}{\sim} u$, then $w \overset{\to}{\nsim} u$.

LEMMA 7.14. *Let c_{frac} be the constant from Lemma 7.13 and consider the case $d < 10^4$. There exists $\varrho_0 \in (0, 1/4)$ and $\epsilon \in (0, 10^{-4}]$ such that if $C_{\mathrm{round}} > C_{5.5}(\epsilon)$ then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process with $|R| \wedge |W| < \varrho_0 n$ satisfying*

$$\max_{b,v,a} \sum_{z \in \mathrm{B}} Q(z, b, v, a, \epsilon) \leq \frac{k}{n} + \frac{1}{16},$$

is (α_2, T)-good, for some universal $\alpha_2 > 0$, and $T = C_{\mathrm{round}}(t_{\mathrm{rel}} + t_*(\epsilon) + s_*(\epsilon))$.

PROOF. We inflate the degree so that $\hat{d} = 10^4$. Without loss of generality, suppose $|R| \leq |W|$.

Notice that a white particle will get pinkened during $(T, T + 1)$ if there exists a red particle satisfying statements 1 to 4 from the proof of Lemma 7.12. We choose the set A to be $A(R)$ from Lemma 7.13.

The first part of the proof proceeds similar to the proof of Lemma 7.12. We obtain the bound

$$\mathbb{E}_M[H_T] \geq c_1 c_2 \sum_{a \in A(R)} \sum_{b : a \overset{\to}{\sim} b} \mathrm{P}\big[a \in GN(R)_\theta, b \in I_{[0,T]}(W)\big].$$

At this point, we refer to the proof of Lemma 6.1, and following the same arguments (using Lemma 7.5 in place of Lemma 5.9) arrive at the analogous statement to (57):

$$\mathbb{E}_M[H_T] \geq \hat{c}_3\left(\frac{1}{64}\mathbb{E}\big[\big|A(R) \cap I_{[0,T]}(R)\big|\big] - \epsilon|R|\right), \tag{59}$$

for some $\hat{c}_3 > 0$. Notice that in applying Lemma 7.5 to obtain the above we have made use of the fact that for a fixed choice of $\hat{d}$ there exists a universal constant D such that $d_{\max}^{\mathrm{in}} \leq D\hat{d}$ (i.e., we take ϵ in Lemma 7.5 to be ϵ/D). Now notice that by Lemmas 5.3, 5.5 (with $\varepsilon = 10^{-4}$), 7.2 (used to argue that $\pi(\mathrm{Nice}(R)) \geq 1 - 32D\varrho_0$) and 7.13 we have

$$\begin{aligned}
&\mathbb{E}\big[\big|A(R) \cap I_{[0,T]}(R)\big|\big] \\
&\quad = |R|\mathrm{P}_{\pi_R}\big[X_T \in A(R)\big] = |R| \sum_{u \in A(R)} \mathrm{P}_{\pi_R}[X_T = u] \\
&\quad \geq c_{\mathrm{frac}}|R| \sum_{u \in \mathrm{Nice}(R)} \mathrm{P}_{\pi_R}[X_T = u] = c_{\mathrm{frac}}\mathbb{E}\big[\big|N(R)\big|\big] \\
&\quad \geq c_{\mathrm{frac}}|R|\big(\pi\big(\mathrm{Nice}(R)\big) - 10^{-4}\big) > c_{\mathrm{frac}}|R|\big(1 - 32D\varrho_0 - 10^{-4}\big).
\end{aligned}$$

Combining this with (59) and taking ϵ and ϱ_0 sufficiently small gives the existence of a universal constant α_2 such that $\mathbb{E}_M[H_T] \geq \alpha_2|R|$. $\square$

PROOF OF PROPOSITION 3.3 FOR $d < 10^4$. Let $t \geq t_0 := t_{\mathrm{mix}}^{(\infty)}(n^{-10})$ and ρ_0 and ϵ be the constants from Lemma 7.14 and suppose $|R| \wedge |W| < \varrho_0 n$. If $C_{\mathrm{round}} > C_{5.5}(10^{-4})$, and $T = C_{\mathrm{round}}(t_{\mathrm{rel}} + t_*(\epsilon) + s_*(\epsilon))$, by Lemma 7.14 (which we can apply here as Lemma 5.11 holds even for modified graphs since the number of interactions is unaffected by the addition of edges which never ring), there exists a universal $\alpha_2 > 0$ such that with probability at least $1 - n^{-10}$, M_t is (α_2, T)-good, that is, $\beta(\alpha_2, t_{\mathrm{round}} - 1) \leq n^{-10}$.

On the other hand, if $|R| \wedge |W| \geq \varrho_0 n$ then set $C_* = C_*^0 \vee \frac{10^4}{\log(1/\varrho_0)}$ (with C_*^0 the constant from Lemma 7.12). Then by Lemma 7.12 with $\varrho = \varrho_0$ there exist constants $\alpha_4 > 0$ and C_{round}^0 such that if $C_{\mathrm{round}} > C_{\mathrm{round}}^0$ then with probability 1, M_t is (α_4, T)-good, for $T = C_{\mathrm{round}} t_{\mathrm{rel}}$, that is, $\beta(\alpha_4, t_{\mathrm{round}} - 1) = 0$.

This completes the proof taking $\alpha = \alpha_2 \wedge \alpha_4$. $\square$

8. Mixing for graphs of high degree: Proof of Theorem 1.3.

Recall that in this regime we have $d \gtrsim \log_{n/k} n$. The analogue of Proposition 3.3 (which gives a bound on the probability that a configuration is not (α, t)-good after a burn-in) for proving Theorem 1.3 is the following proposition.

PROPOSITION 8.1. *There exist constants $\alpha, C_{\mathrm{round}}, C_{\mathrm{deg}} > 0$, such that for all n sufficiently large if $d \geq C_{\mathrm{deg}} \log_{n/k} n$ then $\beta(\alpha, C_{\mathrm{round}} t_{\mathrm{rel}}) \leq n^{-10}$.*

PROOF OF THEOREM 1.3. This is identical to the proof of Theorem 1.1 (general mixing bound) in Section 3.3 using Proposition 8.1 in place of Proposition 3.3. $\square$

In order to prove Proposition 8.1, we must control red and black neighbours of red particles (it is not enough to only consider pairs of red and white particles). We make use of the large degree to argue that after a burn-in period of duration $t_{\mathrm{mix}}^{(\infty)}(n^{-10})$, the probability that the number of black neighbours of a vertex v will be unusually high is extremely small. This considerably simplifies the analysis, as there is no longer a need to consider the number of intersections (during the first t_* time units of the round) between a red particle and its black neighbours (at the end of the constant colour phase of the round). This is what allows us to avoid the term $r_*(c_{1.1}) \log(n/\varepsilon)$ in Theorem 1.3.

Indeed, each neighbour of v is occupied by a black particle with probability at most $\frac{k-1}{n} + n^{-10}$. Using negative association, we can bound the probability that vertex v has at least $(\frac{k}{n} + \zeta)d$ black neighbours, for some $\zeta > 0$, by a bound similar to the probability that a Binomial$(d, \frac{k-1}{n} + n^{-10})$ r.v. is at least $(\frac{k}{n} + \zeta)d$. By assumption on d and k, the last probability can be made n^{-13} (for each fixed $\zeta > 0$), provided C_{deg} is taken to be sufficiently large.

LEMMA 8.2. *Let $\zeta \in (0, 1/16]$ and consider the case $d \geq 10^4 \log_{n/k} n$. If $C_{\mathrm{round}} > C_{5.5}(10^{-4})$ and $T = C_{\mathrm{round}} t_{\mathrm{rel}}$, then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process satisfying*

$$\max_{v \in V} \mathrm{P}\left[\sum_{u:u \sim v} \mathbf{1}_{\{u \in \mathrm{B}_T\}} \geq \left(\frac{k}{n} + \zeta\right) d \,|\, \mathrm{B}_0 = B\right] \leq n^{-10}$$

is (α_3, T)-good, for $\alpha_3 > 0$ a universal constant, and all n sufficiently large.

PROOF. This proof is very similar to the proof of Lemma 6.1. We count H_T in the same way and arrive at the bound (from equation (51))

$$\begin{aligned}\mathbb{E}_M[H_T] \geq {}& \frac{1}{4d}\sum_a \mathrm{P}\big[a \in N(R)\big]\big(1 - L\big(\lambda, \theta, d, |R|\big)\big)(d - \theta d)\\ &- \frac{1}{4d}\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}\mathrm{P}\big[b \in I_{[0,T]}(B), a = I_{[0,T]}(v)\big].\end{aligned} \tag{60}$$

Let $E_\zeta(a)$ be the event that vertex a has less than $(k/n + \zeta)d$ neighbours occupied by black particles at time T. Then by the assumption on M, we have that $\mathrm{P}[E_\zeta(a)^\complement] \leq n^{-10}$. Let $N_t(v)$ be the number of neighbours of vertex v occupied by black particles at time t.

Summing over $a \in \mathrm{Nice(R)}$, $b : a \sim b$ and $v \in R$ in the second double sum on the right-hand side of equation (60) gives

$$\begin{aligned}
&\sum_{a,b:a\sim b}\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}\mathrm{P}\big[b \in I_{[0,T]}(\mathrm{B}), a = I_{[0,T]}(v)\big]\\
&\quad= \sum_a\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}\mathbb{E}\Big[\mathbf{1}_{\{a=I_{[0,T]}(v)\}}\sum_{b:a\sim b}\mathbf{1}_{\{b\in I_{[0,T]}(\mathrm{B})\}}\Big]\\
&\quad= \sum_a\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}\mathbb{E}\big[\mathbf{1}_{\{a=I_{[0,T]}(v)\}}N_T(a)\big]\\
&\quad= \sum_a\sum_{v\in R} \mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}\big(\mathbb{E}\big[\mathbf{1}_{\{E_\zeta(a)\}}\mathbf{1}_{\{a=I_{[0,T]}(v)\}}N_T(a)\big]\\
&\qquad + \mathbb{E}\big[\mathbf{1}_{\{E_\zeta(a)^\complement\}}\mathbf{1}_{\{a=I_{[0,T]}(v)\}}N_T(a)\big]\big)\\
&\quad\leq \sum_a\sum_{v\in R}\big(\mathbf{1}_{\{a\in \mathrm{Nice}(R)\}}(k/n+\zeta)d\mathrm{P}\big[a = I_{[0,T]}(v)\big]\\
&\qquad + d\mathrm{P}\big[E_\zeta(a)^\complement \cap \{a = I_{[0,T]}(v)\}\big]\big)\\
&\quad\leq \sum_a \mathrm{P}\big[a\in N(R)\big](k/n+\zeta)d + d\sum_a \mathrm{P}\big[E_\zeta(a)^\complement\big]\\
&\quad\leq \sum_a \mathrm{P}\big[a\in N(R)\big](k/n+\zeta)d + dn^{-9}.
\end{aligned} \tag{61}$$

Combining equations (60) and (61), we have for any $\theta \in (0, 1)$ and $\lambda > 0$,

$$\begin{aligned}
\mathbb{E}_M[H_T] \geq {}& \frac{1}{4d}\sum_a \mathrm{P}\big[a \in N(R)\big]\big(1 - L\big(\lambda, \theta, d, |R|\big)\big)(d - \theta d)\\
&- \frac{1}{4d}\Big(\sum_a \mathrm{P}\big[a \in N(R)\big](k/n+\zeta)d + dn^{-9}\Big)\\
={}& \frac{1}{4}\sum_a \mathrm{P}\big[a \in N(R)\big]\Big\{\big(1 - L\big(\lambda, \theta, d, |R|\big)\big)(1-\theta) - \frac{k}{n} - \zeta\Big\} - \frac{1}{4}n^{-9}.
\end{aligned}$$

Choosing $\lambda = 0.05$, $\theta = \frac{9}{16} - \frac{k}{2n}$ and using the bound $|R|/n \leq \frac{1}{2} - \frac{k}{2n}$, we have precisely as in the paragraph following (56) that

$$\frac{1}{d}\log L\big(\lambda, \theta, d, |R|\big) = -\lambda\theta + \big(e^\lambda - 1\big)\big(1/32 + |R|/n\big) \leq -0.0008,$$

and so since $\zeta \leq 1/16$ and $d \geq 10^4$ we obtain the bound

$$\mathbb{E}_M[H_T] \geq \frac{1}{64}\mathbb{E}\big[\big|N(R)\big|\big] - \frac{1}{4}n^{-9} \geq \frac{1}{64}\mathbb{E}\big[\big|N(R)\big|\big] - \frac{1}{4}n^{-8}|R|.$$

Notice now that $\mathbb{E}[|N(R)|] = |R|\mathrm{P}_{\pi_R}(X_T \in \mathrm{Nice}(R))$, for (X_t) a realisation of RW(G), and so by Lemmas 5.3 and 5.5 we have that, since $C_{\mathrm{round}} > C_{5.5}(10^{-4})$ (and as there is no degree-inflation $d_{\max}^{\mathrm{in}} = \hat{d}$),

$$\mathbb{E}\big[|N(R)|\big] \geq |R|\big(\pi(\mathrm{Nice}(R)) - 10^{-4}\big) \geq |R|\left(1 - \frac{|R|/n}{1/32 + |R|/n} - 2\times 10^{-4}\right)$$
$$\geq |R|\left(\frac{1}{17} - 2\times 10^{-4}\right).$$

Thus we obtain $\mathbb{E}_M[H_T] \geq \alpha_3|R|$, for all n sufficiently large and any $\alpha_3 \leq 0.0008$. $\square$

PROOF OF PROPOSITION 8.1. If $k \leq 10^{-5}n$, we make use of Lemma 8.2 with $\zeta = 1/16$. Recall the definition of $m_{\frac{1}{16},n,k}$ from Corollary 7.7. We have the bound

$$\frac{1}{32} d m_{\frac{1}{16},n,k} \geq \frac{1}{32} C_{\mathrm{deg}} \log n \left(1 - \frac{\log(16e^2)}{\log(10^5)}\right) \geq \frac{1}{64} C_{\mathrm{deg}} \log n$$

and so combining Corollary 7.7 and Lemma 8.2 with $\zeta = 1/16$ we deduce that if $C_{\mathrm{deg}} \geq 1000$ and $C_{\mathrm{round}} > C_{5.5}(10^{-4})$ then $\beta(\alpha_3, C_{\mathrm{round}} t_{\mathrm{rel}}) \leq n^{-10}$ for some universal $\alpha_3 > 0$. On the other hand, if $k > 10^{-5}n$ then we will instead make use of Lemma 8.2 with $\zeta = \frac{1}{4}\times 10^{-5}$. We have the bound (for each $\varepsilon \in (0,1)$) $\frac{1}{2} d\varepsilon m_{\varepsilon,n,k} \geq \frac{1}{4} d\varepsilon^2 \frac{n}{k}(\frac{1}{2} - \frac{\varepsilon n}{k})$ and so with $\varepsilon = \zeta = \frac{1}{4}\times 10^{-5}$ we obtain $\frac{1}{2} d\varepsilon m_{\varepsilon,n,k} \geq 10^{-13} d$ and, therefore, for C_{deg} sufficiently large (e.g., 10^{21}) and $C_{\mathrm{round}} > C_{5.5}(10^{-4})$ we get using Corollary 7.7 with Lemma 8.2 that $\beta(\alpha_3, C_{\mathrm{round}} t_{\mathrm{rel}}) \leq n^{-10}$ for some universal $\alpha_3 > 0$. $\square$

9. Mixing for sublinear number of particles: Proof of Theorem 1.2.

We consider separately two cases depending on the growth rate of k.

9.1. *The case* $\sqrt{n} \leq k \leq n^{\delta}$. We further split into two sub-cases depending on the degree. The first is for $d \geq C_{\mathrm{deg}}/(1-\delta)$ where C_{deg} is the constant (of the same name) from Theorem 1.3. The proof of Theorem 1.2 in this case follows immediately by Theorem 1.3 (recall that $t_{\mathrm{sp}}(n^{-a}) \asymp_a t_{\mathrm{rel}} \log n$).

The second sub-case is for $d < C_{\mathrm{deg}}/(1-\delta)$. In this case, the analogue of Proposition 3.3 (to obtain a bound on the probability that a configuration is not (α, t)-good after a burn-in) for this case is the following.

PROPOSITION 9.1. *There exist constants $\alpha_\delta, C_{\mathrm{deg}}, C_\delta > 0$, such that for all n sufficiently large if $k \leq n^{\delta}$ and $d < C_{\mathrm{deg}}/(1-\delta)$, then $\beta(\alpha_\delta, C_\delta t_{\mathrm{rel}}) \leq n^{-10}$.*

To complete the proof of Theorem 1.2 for this case, we apply the same arguments as in the proof of Theorem 1.1 (general mixing bound) in Section 3.3 using Proposition 9.1 in place of Proposition 3.3.

To prove Proposition 9.1 in the regime where $|R| \wedge |W|$ is small, we need the following lemma to control black particles. The proof of this lemma is omitted as it is similar to the proof of Lemma 8.2. The degree-inflation referred to in the statement is with $\hat{d} = \lceil C_{\mathrm{deg}}/(1-\delta)\rceil$.

LEMMA 9.2 (Proof omitted). *Let $\delta \in (0,1)$ and $\zeta \in (0, 1/16]$ and consider the case $k \leq n^{\delta}$, $d < \frac{C_{\mathrm{deg}}}{1-\delta}$. There exist constants C_δ, ϱ_δ such that if $C_{\mathrm{round}} > C_\delta$ then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process with $|R| \wedge |W| < \varrho_\delta n$ satisfying*

$$\max_{v\in V} \mathrm{P}\left[\sum_{u: v \overset{\rightarrow}{\sim} u} \mathbf{1}_{\{u \in \mathrm{B}_T\}} \geq (k/n + \zeta)\frac{C_{\mathrm{deg}}}{1-\delta} | \mathrm{B}_0 = B\right] \leq n^{-10}$$

is (α_δ, T)*-good, for* $T = C_{\mathrm{round}} t_{\mathrm{rel}}$, $\alpha_\delta > 0$ *a constant depending only on* δ, *and all* n *sufficiently large.*

PROOF OF PROPOSITION 9.1. Suppose (as in the proof of Proposition 3.3) that $t \geq t_0$. Let ϱ_δ be the constant from Lemma 9.2, let $\hat{d}$ equal $\lceil C_{\mathrm{deg}}/(1-\delta)\rceil$, and suppose $|R| \wedge |W| < \varrho_\delta n$. By Corollary 7.7, we have that for any $\varepsilon > 0$, $B \in (V)_{k-1}$, $v \in V$ and $n = n(\varepsilon)$ sufficiently large,

$$\mathrm{P}\biggl[\mathrm{P}\biggl[\sum_{u:u\sim v} \mathbf{1}_{\{u \in \mathrm{B}_{T+t}\}} \geq \biggl(\frac{k}{n} + \varepsilon\biggr)\hat{d} | \mathcal{F}_t\biggr] \geq \exp\biggl(-\frac{(1-\delta)\hat{d}\varepsilon}{4}\log n\biggr) \biggm| \mathrm{B}_0 = B\biggr]$$
$$\leq \exp\biggl(-\frac{(1-\delta)\hat{d}\varepsilon}{4}\log n\biggr).$$

Taking $\varepsilon = 1/16$, we deduce by Lemma 9.2 that there exist constants $C_\delta, \alpha_\delta > 0$ such that if $C_{\mathrm{round}} > C_\delta$ then with probability at least $1 - n^{-10}$, M_t is (α_δ, T)-good, for $T = C_{\mathrm{round}} t_{\mathrm{rel}}$, that is, $\beta(\alpha_\delta, t_{\mathrm{round}} - 1) \leq n^{-10}$.

On the other hand if $|R| \wedge |W| \geq \varrho_\delta n$, then set $C_* = C_*^0 \vee \frac{C_{\mathrm{deg}}}{(1-\delta)\log(1/\varrho_\delta)}$. Then by Lemma 7.12 with $\varrho = \varrho_\delta$ there exist constants $\alpha_4(\delta) > 0$ and $C_{\mathrm{round}}^0(\delta)$ such that if $C_{\mathrm{round}} > C_{\mathrm{round}}^0(\delta)$ then with probability at least $1 - n^{-10}$, M_t is (α_4, T)-good, for $T = C_{\mathrm{round}} t_{\mathrm{rel}}$, that is, $\beta(\alpha_4, t_{\mathrm{round}} - 1) \leq n^{-10}$. □

9.2. *The case* $k < \sqrt{n}$. In this case, we require a different chameleon process. This new version of the chameleon process has rounds of varying duration. To be precise, the duration of each burn-in period is taken to be $t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$ for some absolute constant $\hat{c} \in (0, 1)$ chosen to be as large as possible while satisfying the requirements described in Remark 7.8. Further, if at the beginning of the jth round we have r red particles, the round starts with an $(\alpha, L(r) - 1)$-good configuration, where if $r \wedge (n - |\mathrm{B}_0| - r) \in (2^{i-1}, 2^i]$ then

$$L(r) = L_i := C_{\mathrm{round}}/\Lambda\bigl(C_{\mathrm{profile}} 2^i/n\bigr) + 1, \tag{62}$$

where $\Lambda(\bullet)$ is as in Section 2.2, for some absolute constants $C_{\mathrm{round}}, C_{\mathrm{profile}} > 0$ to be determined later. The constant-colour "relaxation" phase for such a round is of duration $L(r) - 1$, while the pinkening phase is again of unit length. Thus the duration of the jth round is $t_{\mathrm{round}}(j) := L(|\mathrm{R}_{\rho_j}|)$ and so $\hat{\tau}_j := \rho_j + L(|\mathrm{R}_{\rho_j}|)$, where ρ_j and $\hat{\tau}_j$ still denote the beginning and end of the jth round.

At the end of such a round, we follow the same rule as in the constant-round chameleon depinking procedure, apart from the fact that we replace above $p(M_{\rho_i}, t_{\mathrm{round}} - 1)$ by $p(M_{\rho_i}, L(|\mathrm{R}_{\rho_j}|) - 1)$. If after a depinking time we have r red particles, then we start the following round immediately if the current configuration is $(\alpha, L(r) - 1)$-good. Otherwise, we perform a sequence of burn-in periods of duration $t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$ until the end of the first burn-in period after which we have an $(\alpha, L(r) - 1)$-good configuration, where r denotes the number of red particles at the end of this burn-in. Recall the process $\hat{M}_t$ used in the definition of $\beta(\alpha, t)$ in (37). Let $t_1 := t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$ and for $i \leq \lceil \log_2(\frac{n-k+1}{2})\rceil$ define

$$\beta_i(\alpha) := \max_{B,R,W} \sup_{s \geq t_1} \mathrm{P}\bigl[\hat{M}_s \text{ is not } (\alpha, L_i - 1)\text{-good} \mid \hat{M}_0 = (B, R, W)\bigr], \tag{63}$$

where the maximum is taken over all partitions of V into sets $\mathbf{O}(B)$, R, W satisfying that $|R| \wedge |W| \in (2^{i-1}, 2^i]$ and $B \in (V)_j$ for some $j < \sqrt{n}$ satisfying $\{B(i) : i \in [j]\} = \mathbf{O}(B)$. We will show that if $k = |\mathrm{B}_0| + 1 \leq \sqrt{n}$ then for some absolute constant $\alpha, C > 0$, we have that $\max_{i \leq \lceil \log_2(\frac{n-k+1}{2})\rceil} \beta_i \leq n^{-10}$ (Proposition 9.4).

Recall Proposition 3.8 from Section 3. The following proposition serves as its replacement for this setting. In simple words, it asserts that for some absolute constant M, if no additional burn-in periods occurred (other than the initial one, whose duration is $t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$) by time $t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k) + Mt_{\mathrm{sp}}(\frac{1}{4s})$, then for all $s \in [k, n^3]$ the expected fraction of "missing ink" at time $t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k) + Mt_{\mathrm{sp}}(\frac{1}{4s}) \lesssim t_{\mathrm{sp}}(\frac{1}{4s})$ would be at most s^{-1}. This assertion is similar to the treatment of the chameleon process in [37], where $t_{\mathrm{evolving-sets}}$ is used instead of t_{sp}. While it seems that one can derive it from the analysis in [37], we give a different proof, which we believe to be simpler.

PROPOSITION 9.3 (Proof in in Appendix A). *There exists an absolute constant M such that for all $s \in [k, n^3]$, $k \leq \sqrt{n}$ and $(\mathbf{w}, y) \in (V)_k$ with $\mathbf{w} \in (V)_{k-1}$ and $y \in V$, if we write $\hat{t}(s) := t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k) + Mt_{\mathrm{sp}}(\frac{1}{4s})$, then*

$$\widehat{\mathbb{E}}_{(\mathbf{w},y)}\big[1 - \mathrm{ink}_{\hat{t}(s)}/(n-k+1)\big] \leq s^{-1} + Mt_{\mathrm{sp}}\left(\frac{1}{4s}\right)(n-k+1) \max_{i < \lceil \log_2(n-k+1) \rceil} \beta_i(\alpha). \tag{64}$$

The next proposition serves as the replacement of Proposition 3.3.

PROPOSITION 9.4. *There exist constants $\alpha_{\frac{1}{2}}, C_{\mathrm{round}}, C_{\mathrm{profile}} > 0$, such that for all n sufficiently large if $k \leq \sqrt{n}$ then $\max_{i < \lceil \log_2(n-k+1) \rceil} \beta_i(\alpha_{\frac{1}{2}}) \leq n^{-10}$ (recall that the definition of β_i depends on constants $C_{\mathrm{round}}, C_{\mathrm{profile}}$ through the definition of L_i).*

PROOF OF THEOREM 1.2. The result has already been shown for the case of $\sqrt{n} \leq k \leq n^{\delta}$, so it remains to prove the result for $k < \sqrt{n}$.

Using sub-multiplicativity, we have that $t_{\mathrm{mix}}^{\mathrm{EX}(k)}((2n)^{-i}) \leq i t_{\mathrm{mix}}^{\mathrm{EX}(k)}(\frac{1}{4n})$. It follows that it suffices to consider $\varepsilon \in [\frac{1}{4n}, \frac{1}{4k}]$. Combining Propositions 3.7, 9.3 and 9.4 concludes the proof, upon observing that the term

$$Mt_{\mathrm{sp}}\left(\frac{1}{4s}\right)(n-k+1) \max_{i \leq \lceil \log_2(\frac{n-k+1}{2}) \rceil} \beta_i(\alpha)$$

in the right-hand side of (64) is at most $\lesssim n^{-10} \times n \times n^2 \log n$ (using $\max_i \beta_i(\alpha) \leq n^{-10}$ and $t_{\mathrm{sp}}(\varepsilon) \lesssim t_{\mathrm{rel}} \log(n/\varepsilon) \lesssim n^2 \log(n/\varepsilon)$ for $\varepsilon \leq 1/2$, e.g., [4, 35]). □

We require two lemmas for the proof of Proposition 9.4. The first is a modification of Lemma 5.5 suitable for this setting. Instead of using the Poincaré inequality, we use (30) to bound $\|\mathrm{P}_{\pi_S}[X_T \in \bullet] - \pi\|_{2,\pi}^2$.

LEMMA 9.5 (Proof in the Supplementary Material [17], Appendix C.4). *We denote the uniform distribution on S by π_S. For each $\varepsilon \in (0, 1)$, there exist $C_{9.5}(\varepsilon), C_{\mathrm{p}}(\varepsilon) > 1$ such that for all $C_{\mathrm{round}} > C_{9.5}(\varepsilon)$, $C_{\mathrm{profile}} > C_{\mathrm{p}}(\varepsilon)$ and all $S \subset V$ with $2|S| \leq n$,*

$$\mathrm{P}_{\pi_S}\big[X_T \in \mathrm{Nice}(S)\big] \geq \pi\big(\mathrm{Nice}(S)\big) - \varepsilon,$$

for any $T \geq C_{\mathrm{round}}/\Lambda(C_{\mathrm{profile}}|S|/n)$.

We will also use the following lemma. The proof is very similar to the proof of Lemma 8.2 except instead of using Lemma 5.5 to bound $\mathbb{E}[|N(R)|]$ we use Lemma 9.5 which is where the bound on C_{profile} originates.

LEMMA 9.6 (Proof omitted). *Let $\zeta \in (0, 1/16]$ and consider the case $k \le \sqrt{n}$. There exist constants $C_{\frac{1}{2}}$, C_{p}, $\varrho_{\frac{1}{2}}$ such that if $C_{\mathrm{round}} > C_{\frac{1}{2}}$ and $C_{\mathrm{profile}} > C_{\mathrm{p}}$ then any configuration $M = (B, R, \varnothing, W)$ of the chameleon process with either* (i) $|R| \wedge |W| < \varrho_{\frac{1}{2}} n$ *and* $d < 2 \times 10^4$, *or* (ii) $d \ge 2 \times 10^4$; *and satisfying*

$$\max_{v \in V} \mathrm{P}\Big[\sum_{u: v \overset{\to}{\sim} u} \mathbf{1}_{\{u \in \mathrm{B}_T\}} \ge (k/n + \zeta)\hat{d} | \mathrm{B}_0 = B\Big] \le n^{-10}$$

with $\hat{d} = 2 \times 10^4$ in case (i) *and $\hat{d} = d$ in case* (ii), *is $(\alpha_{\frac{1}{2}}, T)$-good, for $T \ge C_{\mathrm{round}}/\Lambda(C_{\mathrm{profile}}|R|/n)$, $\alpha_{\frac{1}{2}} > 0$ a universal constant, and all n sufficiently large.*

PROOF OF PROPOSITION 9.4. Recall the notation $t_1 := t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k)$ and let $t \ge t_1$. Let $\varrho_{\frac{1}{2}}$ be the constant from Lemma 9.6 and suppose either (i) $|R| \wedge |W| < \varrho_{\frac{1}{2}} n$ and $d < 2 \times 10^4$ or (ii) $d \ge 2 \times 10^4$. By Corollary 7.7 and Remark 7.8 following it (and recalling the choice of $\hat{c}$), we have that for any $\varepsilon > 0$, $B \in (V)_{k-1}$, $v \in V$, and $n = n(\varepsilon)$ sufficiently large,

$$\begin{aligned}
&\mathrm{P}\Big[\mathrm{P}\Big[\sum_{u: u \sim v} \mathbf{1}_{\{u \in \mathrm{B}_{T+t}\}} \ge \Big(\frac{k}{n} + \varepsilon\Big)\hat{d} | \mathcal{F}_t\Big] \ge \exp\Big(-\frac{\hat{d}\varepsilon}{2} \log\big(\varepsilon\sqrt{n}/e^2\big)\Big) \,\Big|\, \mathrm{B}_0 = B\Big] \\
&\quad \le \exp\Big(-\frac{\hat{d}\varepsilon}{2} \log\big(\varepsilon\sqrt{n}/e^2\big)\Big)
\end{aligned}$$

(with $\hat{d} = 2 \times 10^4$ in case (i) and $\hat{d} = d$ in case (ii)). Taking $\varepsilon = 1/16$, we deduce by Lemma 9.6 with $|R| \wedge |W| \in (2^{i-1}, 2^i]$ that there exist constants $C_{\frac{1}{2}}$ and C_{p} such that if $C_{\mathrm{round}} > C_{\frac{1}{2}}$ and $C_{\mathrm{profile}} > C_{\mathrm{p}}$ then with probability at least $1 - n^{-10}$, M_t is $(\alpha_{\frac{1}{2}}, T)$-good, for $T = C_{\mathrm{round}}/\Lambda(C_{\mathrm{profile}} 2^i/n)$, that is, $\beta_i(\alpha_{\frac{1}{2}}) \le n^{-10}$.

On the other hand, if $|R| \wedge |W| \ge \varrho_{\frac{1}{2}} n$, then set $C_{*, \frac{1}{2}} = C_*^0 \vee \frac{10^4}{\delta \log(1/\varrho_{\frac{1}{2}})}$. Then by Lemma 7.12 with $\varrho = \varrho_{\frac{1}{2}}$ there exist constants $\hat{\alpha}_{\frac{1}{2}} > 0$, C_{round}^0, $\hat{C}_{\mathrm{p}}$ (chosen so that for any $|R| \ge \varrho_{\frac{1}{2}} n$ we have $\Lambda(\hat{C}_{\mathrm{p}}|R|/n) = 1/t_{\mathrm{rel}}$) such that if $C_{\mathrm{round}} > C_{\mathrm{round}}^0$ and $C_{\mathrm{profile}} > \hat{C}_{\mathrm{p}}$ then with probability at least $1 - n^{-10}$, M_t is $(\hat{\alpha}_{\frac{1}{2}}, T)$-good, for $T = C_{\mathrm{round}}/\Lambda(C_{\mathrm{profile}} 2^i/n)$, that is, $\beta_i(\hat{\alpha}_{\frac{1}{2}}) \le n^{-10}$. □

10. Lower bounds: Proof of Theorem 1.4 and Proposition 1.7. Recall that $\mathrm{P}_F^{\mathrm{EX}(k)}$ is the distribution of the exclusion process with initial set F. We denote by P_μ^t (resp., P_μ) the distribution of X_t (resp., $(X_t)_{t \ge 0}$), given that the initial distribution is μ.

PROOF OF THEOREM 1.4. By the spectral decomposition, $-\mathcal{L}$ has eigenvalues $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_n$. Denote the corresponding orthonormal basis (w.r.t. $\langle \bullet, \bullet \rangle_\pi$) of eigenvectors by $f_1 = \mathbf{1}, f_2, \dots, f_n$. Without loss of generality, we may assume that $\lambda = \lambda_i$, $f = f_i$ and that $t = t(k, \delta, \varepsilon, \lambda) := \frac{1}{2\lambda}(4\delta \log k - \log(16/\varepsilon)) \ge 0$. Consider $B := \{f \ge 0\}$. Without loss of generality, $|B| \ge n/2$ (otherwise consider $-f$). Let $F \in \binom{V}{k}$ be such that $\mathbb{E}_F[|A_t \cap B|] = \max_{J \in \binom{V}{k}} \mathbb{E}_J[|A_t \cap B|]$. Then by negative correlation

$$\begin{aligned}
\operatorname{Var}_{\pi_{\mathrm{EX}(k)}} |A_0 \cap B| &\le \mathbb{E}_{\pi_{\mathrm{EX}(k)}}\big[|A_0 \cap B|\big] = k\pi(B). \\
\operatorname{Var}_F |A_t \cap B| &\le \mathbb{E}_F\big[|A_t \cap B|\big].
\end{aligned} \tag{65}$$

Denote $\sigma^2 := \frac{1}{2}(\mathrm{Var}_{\pi_{\mathrm{EX}(k)}} |A_0 \cap B| + \mathrm{Var}_F |A_t \cap B|)$. By the standard method of distinguishing statistics [31], Proposition 7.12, if $a := |\mathbb{E}_F[|A_t \cap B|] - \mathbb{E}_{\pi_{\mathrm{EX}(k)}}[|A_0 \cap B|]|^2 \geq 4r\sigma^2$, then

$$\|\mathrm{P}_F^{\mathrm{EX}(k)}(A_t \in \bullet) - \pi_{\mathrm{EX}(k)}\|_{\mathrm{TV}} \geq 1 - \frac{1}{1+r}.$$

We will show that $a \geq 4k/\varepsilon$, which means that we can take above $r = 1/\varepsilon$, as

$$k \geq \frac{1}{2}\big(\mathbb{E}_{\pi_{\mathrm{EX}(k)}}\big[|A_0 \cap B|\big] + \mathbb{E}_F\big[|A_t \cap B|\big]\big) \geq \sigma^2,$$

where the first inequality is trivial and the second inequality follows from (65).

If $D \sim \mathrm{Unif}(\{U \subseteq B : U \in \binom{V}{k}\})$, π_B is the uniform distribution on B and $(X_s)_{s \in \mathbb{R}_+}$ is a random walk on the network $(G, (r_e : e \in E))$ then using the maximality of F (first inequality) and the spectral decomposition in the third equality (namely, $1_B = \pi(B) + \sum_{j=2}^n \sum_{b \in B} \pi(b) f_j(b) f_j$)

$$\begin{aligned}
\mathbb{E}_F\big[|A_t \cap B|\big] &\geq \mathbb{E}_D\big[|A_t \cap B|\big] = k\mathrm{P}_{\pi_B}[X_t \in B] = \frac{k}{\pi(B)}\langle P_t 1_B, 1_B\rangle_\pi \\
&= k\pi(B) + \frac{k}{\pi(B)} \sum_{b' \in B} \pi(b') \sum_{j>1} \sum_{b \in B} \pi(b) f_j(b) f_j(b') e^{-\lambda_j t} \\
\left(\text{write } b_j := \sum_{b \in B} \pi(b) f_j(b)\right) \quad &= k\pi(B) + \frac{k}{\pi(B)} \sum_{j>1} b_j^2 e^{-\lambda_j t} \\
&\geq k\pi(B) + \frac{k}{\pi(B)} b_i^2 e^{-\lambda t} = k\pi(B) + \frac{k}{2\pi(B)} \|f\|_1^2 e^{-\lambda t},
\end{aligned}$$

where we used the fact that $f = f_i$ is orthogonal to $f_1 = \mathbf{1}$, and thus $\mathbb{E}_\pi f = 0$ and $\sum_{b \in B} \pi(b) f_i(b) = \mathbb{E}_\pi[f \vee 0] = \|f\|_1/2$. We get that $a \geq k^2 \|f\|_1^4 e^{-2\lambda t}/4$. By the choice $t = \frac{1}{2\lambda}(4\delta \log k - \log(16/\varepsilon))$ and the assumption $\|f\|_1^4 \geq k^{-1+4\delta}$, we get that $a \geq k^2\|f\|_1^4 e^{-2\lambda t}/4 \geq k(4/\varepsilon) \geq 4\sigma^2/\varepsilon$, as desired. $\square$

REMARK 10.1. It is interesting to note that when $\|f\|_1 \leq k^{-1/8}$ for some unit eigenfunction f as above, it follows from Hölder's inequality that $\|f\|_\infty \geq \|f\|_2^2/\|f\|_1 \geq k^{1/8}$ (the exponent $1/8$ in $\|f\|_1 \leq k^{-1/8}$ is taken as some arbitrary constant smaller than $1/4$, the exponent appearing in Theorem 1.4). In this case, Wilson's method ([42]; see [31], Section 13.5, for a systematic presentation of the method) can sometimes yield that $t_{\mathrm{mix}}^{\mathrm{RW}(1)} \geq c\lambda^{-1}\log k$. We note that in [42] Wilson applied his method to prove a lower bound on the mixing time of $\mathrm{EX}(2^{d-1})$ and $\mathrm{IP}(2^d)$ for the hypercube $\{\pm 1\}^d$. Our argument is different, in that we obtain control on the variances "for free" as a consequence of negative correlation.

PROOF OF PROPOSITION 1.7. As processes $\mathrm{EX}(k)$ and $\mathrm{EX}(n-k)$ are identical it suffices to consider $k \leq n/2$. For fixed such k and $x \in V$, let B be the set of vertices y having the k smallest $P_t(x,y)$ values, where $t = 22t_{\mathrm{mix}}^{\mathrm{EX}(k)}$. By sub-multiplicativity $t \geq 2t_{\mathrm{mix}}^{\mathrm{EX}(k)}(2^{-11})$. Set $\delta := \max_{y \in B} P_t(x,y)$ so that $P_t(x,B) \leq \delta k$. Using the general fact that

$$\min_{C \in \binom{V}{k}} P_t^{\mathrm{EX}(k)}(B,C)/\pi_{\mathrm{EX}(k)}(C) \geq \big(1 - 2\|P_{t/2}^{\mathrm{EX}(k)}(B,\cdot) - \pi_{\mathrm{EX}(k)}\|_{\mathrm{TV}}\big)^2$$

(e.g., Lemma 7 of Chapter 4 of [1]) we also have

$$P_t(x,B) = \sum_{C: x \in C} P_t^{\mathrm{EX}(k)}(B,C) \geq \sum_{C: x \in C} \frac{(1-2^{-10})^2}{\binom{n}{k}} = \frac{k}{n}\big(1-2^{-10}\big)^2.$$

Hence we obtain the bound $\delta \geq \frac{1}{n}(1-2^{-10})^2$. We distinguish between two cases depending on the value of k.

Consider first the case $k \leq n/8$. We have the bound

$$\begin{aligned}\sum_y \left(\frac{1}{n} - P_t(x,y)\right)_+ &= \sum_{y\in B}\left(\frac{1}{n} - P_t(x,y)\right)_+ + \sum_{y\notin B}\left(\frac{1}{n} - P_t(x,y)\right)_+ \\ &\leq \frac{k}{n} + (n-k)\left(\frac{1}{n} - \delta\right) \\ &\leq 1 - (1-2^{-10})^2(1-k/n) < 1/4.\end{aligned}$$

Now consider the case $k > n/8$. We note that $\delta(n-k) \leq \sum_{y\notin B} P_t(x,y) \leq 1$ and so as $k \leq n/2$ we obtain $\delta \leq 2/n$. Thus by a simple counting argument (using $P_t(x,B) \geq \frac{1}{8}(1-2^{-10})^2$ as $k > n/8$) there must be at least $\frac{n}{32}$ vertices $b \in B$ satisfying $P_t(x,b) \geq \frac{1}{8n}(1-2^{-10})^2$. Thus we have the bound

$$\begin{aligned}\sum_y \left(\frac{1}{n} - P_t(x,y)\right)_+ &= \sum_{y\in B}\left(\frac{1}{n} - P_t(x,y)\right)_+ + \sum_{y\notin B}\left(\frac{1}{n} - P_t(x,y)\right)_+ \\ &\leq \frac{n}{32}\left(\frac{1}{n} - \frac{1}{8n}(1-2^{-10})^2\right) + \frac{1}{n}\left(k - \frac{n}{32}\right) + (n-k)\left(\frac{1}{n} - \delta\right) \\ &\leq 1 - (1-2^{-10})^2\left(\frac{1}{2} + \frac{1}{256}\right) < \frac{1}{2} - 2^{-9}.\end{aligned}$$

Thus, combining the two cases, we obtain that for all $k \leq n/2$,

$$t_{\mathrm{mix}}^{\mathrm{RW}(1)}\left(\frac{1}{2} - 2^{-9}\right) \leq 22 t_{\mathrm{mix}}^{\mathrm{EX}(k)},$$

and hence there exists a constant c (2^{-13} suffices) so that $ct_{\mathrm{mix}}^{\mathrm{RW}(1)} \leq t_{\mathrm{mix}}^{\mathrm{EX}(k)}$ for all $k \in [n-1]$. $\square$

11. Examples. We present three additional applications of our results. In the following, we denote by B_r a ball of radius r. We prove the claimed bounds on $r_*(c_{1.1})$ in Section 11.1. In Section 11.2, we show how we can apply our results to product graphs and in particular obtain the bounds claimed in Section 1.1 for the hypercube.

(i) For an n-vertex d-regular vertex-transitive graph satisfying $|B_r| \geq ce^{cr}$ for all r such that $|B_r| \leq \frac{4}{c_{1.1}} \log n$, for some $c > 0$, we have that $r_*(c_{1.1}) \lesssim d^2(\log\log n)^3$ (see Proposition 11.3). Hence (by Theorem 1.1) $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim t_{\mathrm{rel}} \log n$, provided that $t_{\mathrm{rel}} \gtrsim d^2(\log\log n)^3$.

(ii) For an n-vertex d-regular vertex-transitive graph satisfying $|B_r| \geq ce^{cr^\alpha}$ for all r such that $|B_r| \leq \frac{4}{c_{1.1}} \log n$, for some $\alpha \in (0,1)$ and $c > 0$, we have that $r_*(c_{1.1}) \lesssim d^2(\log\log n)^{1+\frac{2}{\alpha}}$ (see Proposition 11.3). Hence $\max_k t_{\mathrm{mix}}^{\mathrm{EX}(k)} \lesssim t_{\mathrm{rel}} \log n$, provided that $t_{\mathrm{rel}} \gtrsim d^2(\log\log n)^{1+\frac{2}{\alpha}}$. In particular, this holds if $|B_r| \leq Ce^{Cr^\beta}$ for all r, for some $\beta \in (0,1)$ and $C > 0$ (as this implies that $t_{\mathrm{rel}} \gtrsim \frac{\mathrm{Diameter}}{\log n} \gtrsim (\log n)^{(1-\beta)/\beta}$).

(iii) The following example is taken from [14], Section 4.2.1 (we refer the reader there for the relevant definitions; see also [13], where it is shown that Cayley graphs of moderate growth satisfy a local-Poincaré inequality, and many other examples are given).

If G is a d-regular graph of diameter γ and (A,c)-moderate growth, satisfying a local-Poincaré inequality with a constant a, then $t_{\mathrm{rel}} \asymp \gamma^2 \asymp t_{\mathrm{sp}}(\frac{1}{4})$ (with the implicit constants depending on d, A, c and a); $t_{\mathrm{rel}} \asymp \gamma^2$ is due to Diaconis and Saloff-Coste [11], Theorem 3.1 (cf. our Secction 11.1). By Corollary 1.8, $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp_{a,d,A,c} \gamma^2 \log(k+1)$ uniformly in k.

11.1. *Vertex-transitive graphs and the giant component of super-critical percolation.* Let $G = (V, E)$ be an n-vertex connected graph. We say that G is *vertex-transitive* if the action of its automorphism group on its vertices is transitive. Denote the volume of a ball of radius r in G by $V(r)$. Denote the diameter of G by $\gamma := \inf\{r : V(r) \geq n\}$. Following Diaconis and Saloff-Coste, we say that G has (c, a)-*moderate growth* if $V(r) \geq cn(r/\gamma)^a$. Breuillard and Tointon [6] proved that for Cayley graphs of constant degree, this condition is equivalent in some quantitative sense to the condition that $n \leq \beta\gamma^\alpha$ for some $\alpha, \beta > 0$, which is of course a much simpler condition. Let P be the transition matrix of simple random walk (SRW) on G. We consider the case of continuous-time SRW with $\mathcal{L} = P - I$. Diaconis and Saloff-Coste [11] showed that for a Cayley graph G of (c, a)-moderate growth we have

$$c^2\gamma^2 4^{-2a-1} \lesssim t_{\mathrm{rel}} \lesssim t_{\mathrm{mix}}^{(\infty)} \lesssim_{c,a} \gamma^2.$$

We note that the proof of $c^2\gamma^2 4^{-2a-1} \lesssim t_{\mathrm{rel}}$ works even if G is merely vertex-transitive of (c, a)-moderate growth. Namely, they argue that the function $h(x) := \mathrm{distance}(x, \mathbf{o})$ (for some arbitrary $\mathbf{o} \in V$) satisfies that $\mathrm{Var}_\pi h/\mathcal{E}(h, h) \geq \mathrm{Var}_\pi h \geq \gamma^2(V(\lfloor\gamma/4\rfloor)/2n)^2 \geq c^2\gamma^2 4^{-2a-1}$. Indeed, if $h(x) = \gamma$ then for the vertices y in the ball of radius $r := \lfloor\gamma/4\rfloor$ centered at x (resp., $\mathbf{o}$) we have $h(y) \geq \frac{3}{4}\gamma$ (resp., $\leq \frac{\gamma}{4}$). Denote these two balls by $B_x(r)$ and $B_{\mathbf{o}}(r)$. If X, Y are i.i.d. $\pi = \mathrm{Unif}(V)$, then

$$\mathrm{Var}_\pi h = \frac{1}{2}\mathbb{E}\big[\big(h(X) - h(Y)\big)^2\big] \geq \frac{\gamma^2}{8}\pi\big(B_x(r)\big)\pi\big(B_{\mathbf{o}}(r)\big). \tag{66}$$

Lyons et al. [34], Lemma 7.2, showed that for an n-vertex vertex-transitive graph, for all $A \subset V$ such that $|A| \leq n/2$ we have

$$\frac{|\partial_{\mathrm{V}}^{\mathrm{in}} A|}{|A|} \geq \frac{1}{2R(2|A|)}, \tag{67}$$

where $\partial_{\mathrm{V}}^{\mathrm{in}} A := \{a \in A : P(a, A^c) > 0\}$ is the internal vertex boundary of A and $R(m) := \inf\{r : V(r) \geq m\}$ is the inverse growth function (note that $R(m) = \infty$ for $m > n$, and so (67) holds trivially when $|A| > n/2$). (Lemma 7.2 in [34] is stated for infinite unimodular graphs, but the proof works verbatim for finite transitive graphs, which are always unimodular. See also [36], Lemma 10.46, where G is not assumed to be infinite.)

PROPOSITION 11.1. *If G is a d-regular vertex-transitive of (c, a)-moderate growth, then*

$$c^2\gamma^2 4^{-2a-1} \leq t_{\mathrm{rel}} \lesssim t_{\mathrm{evolving-sets}}(1/4) \lesssim a(2/c)^{2/a} d^2\gamma^2. \tag{68}$$

Consequentially, (uniformly in k)

$$t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp_{c,a,d} \gamma^2 \log(k+1). \tag{69}$$

Similarly, if G is the largest connected component of super-critical percolation on $(\mathbb{Z}/L\mathbb{Z})^d$ with parameter p then w.h.p. (*as $L \to \infty$*)

$$\gamma^2 \lesssim_{d,p} t_{\mathrm{rel}} \lesssim t_{\mathrm{evolving-sets}}(1/4) \lesssim_{d,p} \gamma^2. \tag{70}$$

Consequentially, w.h.p. (*uniformly in k*)

$$t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp_{d,p} \gamma^2 \log(k+1). \tag{71}$$

REMARK 11.2. In the setup of (68), the bound obtained on $t_{\mathrm{mix}}^{(\infty)}$ in [35] via the spectral measure is often better than the one obtained via $t_{\mathrm{evolving-sets}}$.

PROOF. We first note that (69) and (71) follow by combining (68) and (70) with (10). The first inequality in (68) was discussed above. The corresponding bound in (70) is obtained by considering the same h as in (66) (noting that the size of any ball of radius $\lfloor \gamma/4 \rfloor$ in the giant component w.h.p. has volume comparable to the total number of vertices). The middle inequality in (68) and (70) follows from (28) and (19). The last inequality in (70) is taken from [40].

The proof of the last inequality in (68) follows by plugging in (26) the estimate $\Phi^{-2}(\delta) \lesssim d^2\gamma^2(\frac{2\delta}{c})^{2/a}$, which can be derived via (67). $\square$

PROPOSITION 11.3. *If G is a d-regular vertex-transitive graph of size n as in Example* (i) (*resp.*, (ii)) *then*

$$r_*(c_{1.1}) \le Cd^2(\log\log n)^3 \quad \left(\textit{resp. } r_*(c_{1.1}) \le Cd^2(\log\log n)^{1+\frac{2}{\alpha}}\right).$$

PROOF. As above, use (67) to bound $\Phi^{-2}(\delta)$ for all $\delta \le 4(\log n)/(c_{1.1}n)$. In the setup of Example (i) (67) yields that $\Phi^{-2}(\delta) \lesssim [d\log(\delta n)]^2$ and in that of Example (ii) that $\Phi^{-2}(\delta) \lesssim d^2[\log(n\delta)]^{2/\alpha}$. The assertion of the proposition now follows from (28) with $\varepsilon = c_{1.1}n/\log n$. $\square$

11.2. *The hypercube and product graphs.* We now consider the hypercube $\{\pm 1\}^d$. We consider the case that each edge has rate $1/d$. Then $t_{\mathrm{rel}} = \frac{d}{2} = 2/c_{\mathrm{LS}}$ (see [12]). By Proposition 2.4, it is easy to verify that $t_{\mathrm{sp}}(\frac{1}{2}) \lesssim d\log d \asymp t_{\mathrm{mix}}$. By (31) $r_*(c_{1.1}) \lesssim \log d \ll t_{\mathrm{rel}}$. By Theorem 1.1 in conjunction with Corollaries 1.9 and 1.11, we get that $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp d\log(dk)$ (for $k \le d$, we use Theorem 1.4 to argue that $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \gtrsim d\log d$).

DEFINITION 11.4. The Cartesian product $G_1 \times G_2 = (V', E')$ of two graphs $G_i = (V_i, E_i)$ is defined via $V' := V_1 \times V_2$ and $E' := \{\{(v_1, v_2), (u_1, u_2)\} : v_1 = u_1 \in V_1 \text{ and } u_1u_2 \in E_2, \text{ or vice versa}\}$. For a graph $G = (V, E)$, we denote the d-fold self (Cartesian) product of G with itself by $G_{\otimes d} = (V^n, E(G_{\otimes n}))$, that is, $G_{\otimes d} = G_{\otimes(d-1)} \times G = G \times \cdots \times G$.

Note that the d-dim hypercube is the d-fold self-product of the complete graph on two vertices with itself. Consider the case that G is d_G-regular. Then $G_{\otimes d}$ is $d \times d_G$ regular. Consider EX(k) on $G_{\otimes d}$ in which each edge rings at rate $\frac{1}{dd_G}$. We extend the above argument and show that $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp d\log(dk)$ (all asymptotic notation here is as $d \to \infty$ for a fixed G and some implicit constants may depend on G). We note that the analysis below can easily be extended to the case of EX(k) on $G_1 \times \cdots \times G_d$ with the G_i's being of uniformly bounded size. The regularity condition can be lifted as well. Indeed if the G_i's are of uniformly bounded size, then all vertices of $G_1 \times \cdots \times G_d$ are of degree proportional to d. As explained in the Supplementary Material [17], Appendix B, our analysis can be extended to cover this case.

A general result about the log-Sobolev constant of a product chain asserts that $c_{\mathrm{LS}}(G_{\otimes d})$ the log-Sobolev constant of the random walk on $G_{\otimes d}$ (with the above rates) is $c_{\mathrm{LS}}(G)/d$, where $c_{\mathrm{LS}}(G)$ is the log-Sobolev constant of a random walk on G, and likewise $t_{\mathrm{rel}}(G_{\otimes d}) = dt_{\mathrm{rel}}(G)$ (see [12]). By Proposition 2.4 we get that $t_{\mathrm{sp}}^{G_{\otimes d}}(\frac{1}{2}) \lesssim \frac{d\log(d\log|G|)}{c_{\mathrm{LS}}(G)}$. By (31), $r_*(c_{1.1}) \lesssim \frac{\log(d\log|G|)}{c_{\mathrm{LS}}(G)\log|G|} \ll t_{\mathrm{rel}}(G_{\otimes d})$. Finally, we have that $t_{\mathrm{mix}}(G_{\otimes d}) = \frac{t_{\mathrm{rel}}(G)}{2}d\log d(1 \pm o(1))$ [31], Theorem 20.7. In particular, we have that $t_{\mathrm{sp}}^{G_{\otimes d}}(\frac{1}{2}) \lesssim t_{\mathrm{mix}}(G_{\otimes d})$. As before, by Theorem 1.1 in conjunction with Corollaries 1.9 and 1.11 we get that $t_{\mathrm{mix}}^{\mathrm{EX}(k)} \asymp d\log(dk)$, as claimed.

APPENDIX A: PROOF OF PROPOSITION 7.17

We consider the case $k \le \sqrt{n}$, where the duration of a round of the chameleon process, starting with r red particles such that $r \wedge (n-k+1-r) \in (2^{i-1}, 2^i]$ is $L(r) = L_i$ as defined in (62). By (27) and the fact that $\Lambda(\varepsilon)$ is nondecreasing in ε, we obtain the following.

LEMMA A.1. *For all $\varepsilon \in (0, 1/2)$, we have that $\Lambda(\varepsilon) \le -2\min_x \mathcal{L}(x,x)$. In particular, in our setup $\Lambda(\varepsilon) \le 2$ for all ε and so $L_i \le (C_{\text{round}} + 2)/\Lambda(C_{\text{profile}} 2^i/n)$.*

While we are really interested in studying the process $(\text{ink}_t)_{t \ge 0}$ (conditioned on Fill), it is more convenient to study the related process $(\widehat{Y}_t)_{t \in \mathbb{R}_+}$ on $[n-k+1]$ which is defined by the following rule. Whenever it reaches state r, it stays put for $L(r)$ time units before making a step according to $\hat{P}$, the transition matrix of $\mathbf{Y} := (Y_i)_{i \in \mathbb{Z}_+}$ (defined in the Supplementary Material [17], Appendix C.3).

Recall that in this setup, each burn-in period has duration $t_{\text{mix}}^{(\infty)}(\hat{c}/k)$, where $\hat{c}$ is some absolute constant (and again, the process starts with an initial burn-in period). Let BIP be the set of all times which are part of a burn-in period of the chameleon process. For all $s \ge 0$, let $t(s) := \inf\{t \notin \text{BIP} : t - j(t) t_{\text{mix}}^{(\infty)}(\hat{c}/k) = s\}$, where $j(t)$ is the number of burn-in periods by time t. Then $(\widehat{Y}_s)_{s \in \mathbb{R}_+}$ has the same distribution as that of $(\text{ink}_{t(s)})_{s \in \mathbb{R}_+}$ conditioned on Fill. Since typically $s - t(s) \ll s$, we may translate estimates concerning $(\widehat{Y}_s)_{s \in \mathbb{R}_+}$ to ones concerning $(\text{ink}_t)_{t \ge 0}$. Before diving into the analysis of $\widehat{\mathbf{Y}} := (\widehat{Y}_t)_{t \ge 0}$, we need the following simple proposition concerning $\mathbf{Y}$.

Let $\hat{\ell} := \lceil \log_2(n-k+1) \rceil - 1$ and $m := \lceil (n-k+1)/2 \rceil$. Our strategy is to decompose the process ink_t given Fill into three stages: (1) The time until it hits $[m-1]^{\complement}$, (2) the additional time from that moment until it never goes below m, and (3) the remaining time. The idea is that the process viewed at stage (3) is like $(\text{ink}_t : t \ge 0)$ started above $m-1$, conditioned on hitting $n-k+1$ before $[m-1]$. A similar supermartingale as in [17], Lemma C.1, can be used to study this process, with the crucial key difference that now we do not pick up a factor of $\sqrt{n-k+1}$ (as now $I_i \ge \frac{1}{2}$). It remains to find bounds t_i such that the probability that the duration of stage $i \in \{1,2\}$ is more that t_i is $o(\varepsilon/k)$. This is done by first showing that for the chain $\mathbf{Y}$ various relevant quantities have uniform exponential tails, and then translating this into corresponding statements about $\widehat{\mathbf{Y}}$.

For $i \le \hat{\ell}$, let

$$
\begin{aligned}
T_i^{\uparrow} &:= \inf\{j : Y_j \ge 2^i \wedge m\} = \text{The hitting time of } \big[(2^i \wedge m) - 1\big]^{\complement} \\
T_{[m-1]} &:= \inf\{j : Y_j < m\} = \text{The hitting time of } [m-1], \\
S &:= \inf\Big\{j : \min_{s: s \ge j} Y_s \ge m\Big\} - T_{\hat{\ell}}^{\uparrow} = \text{Time between the first visit to} \\
&\qquad [m-1]^{\complement} \text{ and the time following the last visit to } [m-1], \\
\text{Cross} &:= \big|\{i : Y_{i+1} < m \le Y_i\}\big| = \text{number of down-crossings below } m.
\end{aligned}
\tag{72}
$$

PROPOSITION A.2. *There exist absolute constants $0 < c_i < 1 < C_i$ (for $i \in [6]$) such that*:

(i)

$$\forall s, \quad \max_{i \le \hat{\ell}} \max_{r \in [2^{i-1}, 2^i)} \mathrm{P}_r\big[T_i^{\uparrow} > s\big] \le C_1 \exp(-c_1 s). \tag{73}$$

Hence for some $c_6 \in (0, c_1/2)$, for all $\gamma \in (0, c_6)$ we have that

$$\max_{i \le \hat{\ell}} \max_{r \in [2^{i-1}, 2^i)} \mathbb{E}_r\big[\exp\big(\gamma T_i^{\uparrow}\big)\big] \le \exp(C_6 \gamma). \tag{74}$$

(ii) *Let* $I := [m, \frac{3}{2}m]$. *Then*

$$\forall s, \quad \max_{r \in I} \mathrm{P}_r[\mathrm{Cross} > s] \le C_2 \exp(-c_2 s). \tag{75}$$

$$\forall s, \quad \max_{r \in I} \mathrm{P}_r[T_{[m-1]} \mid \mathrm{Cross} \ge 1] \le C_3 \exp(-c_3 s). \tag{76}$$

(iii)

$$\forall s, \quad \max_{r \in I} \mathrm{P}_r[S \ge s] \le C_4 \exp(-c_4 s). \tag{77}$$

(iv) *For all* $r \in I$, *conditioned on* $Y_0 = r$ *and* $S = 0$ *we have that* $c_5^{-i}(1 - \frac{Y_i}{n-k+1})$ *is a supermartingale* ($c_5 = c_5(\alpha, p)$, *where* α *is as in the definition of* $\Delta(r)$).

PROOF. We first prove (73). Let $U_t := |\{j \le t : Y_j > Y_{j-1}\}|$ and $D_t = |\{j \le t : Y_j < Y_{j-1}\}|$. Up to a rounding error (resulting from the ceiling in the definition of $\Delta(r)$), whenever the size of Y_i changes, it is multiplied by a factor of either $1 + \alpha$ or $1 - \alpha$. Using the fact that $(1+\alpha)^{1+\alpha}(1-\alpha)^{1-\alpha} > 1$ for all $\alpha \in (0, 1)$ (and so also $(1+\alpha)^{p\frac{1+\alpha}{2}}(1-\alpha)^{p\frac{1-\alpha}{2}} > 1$), ignoring the rounding error we get that there exists some $\varepsilon > 0$ and C_ε such that for all $i \le \hat{\ell}$ and all $r \in [2^{i-1}, 2^i)$, if $s \ge C_\epsilon$, $U_s \ge ps(\frac{1+\alpha}{2} - \varepsilon)$ and $D_s > ps(\frac{1-\alpha}{2} + \varepsilon)$ then $T_i^{\uparrow} \le s$. It is easy to verify that this implies (73), as the probability that this fails for some fixed s decays exponentially in s (uniformly). To deal with the rounding error, one can control its possible effects whenever Y_i is at least some constant $C \in \mathbb{N}$. Thus by the above reasoning $\max_{i \le \hat{\ell}} \max_{r \in [2^{i-1}, 2^i)} \mathrm{P}_r[|\{t \le T_i^{\uparrow} : Y_t \ge C\}| > s] \le C' e^{-ct}$ for all s. Hence, it suffices to argue that $\max_{i \le \hat{\ell}} \max_{r \in [2^{i-1}, 2^i)} \mathrm{P}_r[|\{t \le T_i^{\uparrow} : Y_t < C\}| > s] \le C' e^{-ct}$ for all s. This follows from the fact that

$$\max_{i \le \hat{\ell}} \max_{r \in [2^{i-1}, 2^i)} \mathrm{P}_r\big[\big|\{t \le T_i^{\uparrow} : Y_{t+1} < C \le Y_t\}\big| > s\big] \le C' e^{-ct}$$

for all s. We leave the details as an exercise.

Observe that (74) follows easily from (73). We now prove (77). It suffices to show that $\max_{r \in I} \mathbb{E}[z^S] < \infty$ for some $z > 1$. We may write $S = \sum_{i=1}^{\mathrm{Cross}} K_i$, where K_i is the time the chain spends above m during its ith epoch above m. Noting that by part (ii) $M(z) := \max_{r' \in I} \mathbb{E}_{r'}[z^{K_1}]$ satisfies $\lim_{z \to 1^+} M(z) = 1$, and $\mathbb{E}[z^{\mathrm{Cross}}] < \infty$ for all $0 < z \le z_0 > 1$. As $\alpha \in (0, 1/2)$, it follows that if $Y_i < m < Y_{i+1}$ then $Y_{i+1} \in I$. Hence by the strong Markov property, for some $z > 1$,

$$\max_{r \in I} \mathbb{E}[z^S] \le \max_{r \in I} \mathbb{E}_r[M(z)^{\mathrm{Cross}}] < \infty.$$

The proof of part (iv) is analogous to that of [17], Lemma C.1, and is thus omitted.

Inequality (75) follows from the fact that for every fixed $\varepsilon > 0$ with positive probability we have that $U_s \ge \lceil ps(\frac{1+\alpha}{2} - \varepsilon)\rceil$ and $D_s > \lfloor ps(\frac{1-\alpha}{2} + \varepsilon)\rfloor$ for all $s > 0$, and this probability is uniform in $r \in n - k + 1$. Thus $a_* := \min_{r \ge m} \mathrm{P}_x[\mathrm{Cross} = 0]$ is bounded from below (uniformly in $n - k + 1$) and by the strong Markov property Cross is stochastically dominated by the (shifted) Geometric distribution of parameter a_*.

Finally, (76) follows by considering the Doob's transform of $\mathbf{Y}$ obtained by conditioning on $T_{[m-1]} < \infty$. An elementary calculation shows that under this conditioning, up to time $T_{[m-1]} < \infty$ the chain has transition probabilities Q satisfying $Q(r, r - \Delta(r)) < Q(r, r + \Delta(r))$ while for $r \in I' := \{\frac{3}{2}m, \dots, n - k + 1 - 1\}$ we have $Q(r, r - \Delta(r)) < c'_{\alpha,p} Q(r, r + \Delta(r))$ for some $c'_{\alpha,p} \in (0, 1)$ (independent of $n - k + 1$). We may write $T_{[m-1]} := \sum_{j=1}^{\widehat{\mathrm{Cross}}} F_j + F'_j$, where $\widehat{\mathrm{Cross}}$ is the number of times the chain enters the interval I' and then leaves it, F_i

(resp., F_i') is the time it spends in I (resp., I') during the ith epoch. As above, it is not hard to verify that $\widehat{\mathrm{Cross}}$, the F_i's and the F_i''s have uniformly exponentially decaying tails. This implies the assertion of part (iv) in a similar fashion to the derivation of part (iii) from part (ii). We leave the details as an exercise. □

PROPOSITION A.3. *Let* $\tau := \inf\{t : \min_{s:s\geq t} \widehat{Y}_s \geq m\}$. *Then* (*starting from* $\widehat{Y}_0 = 0$)

$$\mathbb{E}\left[1 - \frac{\widehat{Y}_{s+t}}{n-k+1}\right] \leq \mathrm{P}[\tau \geq t] + C\exp(-cs/t_{\mathrm{rel}}), \tag{78}$$

$$\mathrm{P}\left[\tau \geq Ct_{\mathrm{sp}}\left(\frac{\varepsilon}{4k}\right)\right] \leq \frac{\varepsilon}{16k^2}. \tag{79}$$

PROOF. Observe that (78) is a direct consequence of part (iv) of Proposition A.2. We now prove (79). We use the same notation as in (72), but now for the chain $\widehat{\mathbf{Y}}$. In this notation, $\tau = S + \sum_{i\in[\hat{\ell}]} T_i^{\uparrow}$. By (77), for all $s \geq 0$, $\mathrm{P}[S \geq st_{\mathrm{rel}}] \leq C_4\exp(-c_4 s)$. Hence

$$\mathrm{P}\big[S \geq C' t_{\mathrm{rel}} \log(k/\varepsilon)\big] \leq \frac{\varepsilon}{32k^2}. \tag{80}$$

By (74), there exist $c \in (0, 1)$ and C_6 such that for all $\gamma \leq c/t_{\mathrm{rel}}$ and all $i \leq \hat{\ell}$ we have

$$\max_{r\in[2^{i-1},2^i)} \mathbb{E}_r\big[\exp(\gamma T_i^{\uparrow})\big] \leq \exp(C_6\gamma L_i), \tag{81}$$

where $L_i \leq (C_{\mathrm{round}} + 2)/\Lambda(C_{\mathrm{profile}} 2^i/n)$ by Lemma A.1. Thus,

$$\begin{aligned}\mathbb{E}\big[e^{\gamma(\tau-S)}\big] = \mathbb{E}\left[\exp\left(\gamma \sum_{i\in[\hat{\ell}]} T_i^{\uparrow}\right)\right] &\leq \exp\left(C_6\gamma \sum_{i\in[\hat{\ell}]} L_i\right) \\ &\leq \exp\left(C_7\gamma t_{\mathrm{sp}}\left(\frac{1}{4k}\right)\right).\end{aligned}$$

Picking $C_8 = 6(C_7 \vee 1)/c$ and $\gamma = c/t_{\mathrm{rel}}$, we get that

$$\begin{aligned}\mathrm{P}\left[\tau - S \geq C_8 t_{\mathrm{sp}}\left(\frac{\varepsilon}{4k}\right)\right] &\leq \mathbb{E}\big[e^{\gamma(\tau-S)}\big]e^{-\gamma C_8 t_{\mathrm{sp}}(\frac{\varepsilon}{4k})} \\ &\leq e^{5(C_7\vee 1)t_{\mathrm{sp}}(\frac{\varepsilon}{4k})/t_{\mathrm{rel}}} \leq \frac{\varepsilon}{32k^2},\end{aligned}$$

where we have used the fact that $t_{\mathrm{sp}}(\frac{\varepsilon}{4k}) \geq t_{\mathrm{rel}} \log(\frac{2k}{\varepsilon})$. This, in conjunction with (80), concludes the proof. □

PROOF OF PROPOSITION 9.3. Let $s \in [k, n^3]$. Let $M \geq 1$ be some absolute constant to be determined shortly. Recall that $\hat{t}(s) := t_{\mathrm{mix}}^{(\infty)}(\hat{c}/k) + q$, where $q = q(s, M) := Mt_{\mathrm{sp}}(\frac{1}{4s})$. By Proposition A.3, we may pick M such that $q \geq C[t_{\mathrm{sp}}(\frac{1}{4s}) + t_{\mathrm{rel}} \log(\frac{1}{4s})]$, and so

$$\begin{aligned}\widehat{\mathbb{E}}_{(\mathbf{w},y)}\big[1 - \mathrm{ink}_{\hat{t}(s)}/(n-k+1)\big] &\leq \mathbb{E}\left[1 - \frac{\widehat{Y}_q}{n-k+1}\right] + \widehat{\mathrm{P}}_{(\mathbf{w},y)}\big[j\big(\hat{t}(s)\big) \geq 2\big] \\ &\leq s^{-1} + \widehat{\mathrm{P}}_{(\mathbf{w},y)}\big[j\big(\hat{t}(s)\big) \geq 2\big].\end{aligned}$$

Finally, $\widehat{\mathrm{P}}_{(\mathbf{w},y)}[j(\hat{t}(s)) \geq 2] \leq (n-k+1)\mathrm{P}_{(\mathbf{w},y)}[j(\hat{t}(s)) \geq 2] \leq (n-k+1)q \max_i \beta_i(\alpha)$, by a simple union bound (over all rounds by time $\hat{t}(s)$), using the fact that the duration of each round is at least 1 time unit. □

Acknowledgments. The authors are grateful to Nathanaël Berestycki, Gady Kozma, Ben Morris and Roberto Oliveira for useful discussions. In particular, we wish to thank Gady Kozma for pointing out that the mixing time in the case of the hypercube was previously considered in [42].

The first author was financially supported by the EPSRC grant EP/L018896/1.

SUPPLEMENTARY MATERIAL

Appendices B and C (DOI: 10.1214/20-AOP1455SUPP; .pdf). In two further Appendices, we provide details on the ways to relax some assumptions as discussed in Section 1.4 and give the technical proofs of Corollaries 1.8–1.11, Proposition 3.4, Proposition 3.7, Proposition 3.8, Lemma 5.5, Lemma 5.11, Lemma 7.6 and Lemma 7.13.

REFERENCES

[1] ALDOUS, D. and FILL, J. (2002). Reversible Markov chains and random walks on graphs.

[2] ALON, G. and KOZMA, G. (2018). Comparing with octopi. Preprint. Available at arXiv:1811.10537.

[3] BENJAMINI, I. and HERMON, J. (2019). Rapid social connectivity. *Electron. J. Probab.* **24** Paper No. 32, 33. MR3940762 https://doi.org/10.1214/19-EJP294

[4] BOCZKOWSKI, L., PERES, Y. and SOUSI, P. (2018). Sensitivity of mixing times in Eulerian digraphs. *SIAM J. Discrete Math.* **32** 624–655. MR3775128 https://doi.org/10.1137/16M1073376

[5] BORCEA, J., BRÄNDÉN, P. and LIGGETT, T. M. (2009). Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* **22** 521–567. MR2476782 https://doi.org/10.1090/S0894-0347-08-00618-8

[6] BREUILLARD, E. and TOINTON, M. C. H. (2016). Nilprogressions and groups with moderate growth. *Adv. Math.* **289** 1008–1055. MR3439705 https://doi.org/10.1016/j.aim.2015.11.025

[7] CANCRINI, N. and MARTINELLI, F. (2000). On the spectral gap of Kawasaki dynamics under a mixing condition revisited *J. Math. Phys.* **41** 1391–1423. MR1757965 https://doi.org/10.1063/1.533192

[8] CAPUTO, P., LIGGETT, T. M. and RICHTHAMMER, T. (2010). Proof of Aldous' spectral gap conjecture. *J. Amer. Math. Soc.* **23** 831–851. MR2629990 https://doi.org/10.1090/S0894-0347-10-00659-4

[9] CONNOR, S. B. and PYMAR, R. J. (2019). Mixing times for exclusion processes on hypergraphs. *Electron. J. Probab.* **24** Paper No. 73, 48. MR3978223 https://doi.org/10.1214/19-EJP332

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

[11] DIACONIS, P. and SALOFF-COSTE, L. (1994). Moderate growth and random walk on finite groups. *Geom. Funct. Anal.* **4** 1–36. MR1254308 https://doi.org/10.1007/BF01898359

[12] DIACONIS, P. and SALOFF-COSTE, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** 695–750. MR1410112 https://doi.org/10.1214/aoap/1034968224

[13] DIACONIS, P. and SALOFF-COSTE, L. (1996). Nash inequalities for finite Markov chains. *J. Theoret. Probab.* **9** 459–510. MR1385408 https://doi.org/10.1007/BF02214660

[14] GOEL, S., MONTENEGRO, R. and TETALI, P. (2006). Mixing time bounds via the spectral profile. *Electron. J. Probab.* **11** 1–26. MR2199053 https://doi.org/10.1214/EJP.v11-300

[15] GREENBERG, S., PASCOE, A. and RANDALL, D. (2009). Sampling biased lattice configurations using exponential metrics. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* 76–85. SIAM, Philadelphia, PA. MR2809307

[16] HERMON, J. and PERES, Y. (2018). A characterization of L_2 mixing and hypercontractivity via hitting times and maximal inequalities. *Probab. Theory Related Fields* **170** 769–800. MR3773799 https://doi.org/10.1007/s00440-017-0769-x

[17] HERMON, J. and PYMAR, R. (2020). Supplement to "The exclusion process mixes (almost) faster than independent particles." https://doi.org/10.1214/20-AOP1455SUPP

[18] HERMON, J. and SALEZ, J. (2019). The interchange process on high-dimensional products. Preprint. Available at arXiv:1905.02146.

[19] HERMON, J. and SALEZ, J. (2019). A version of Aldous' spectral-gap conjecture for the zero range process. *Ann. Appl. Probab.* **29** 2217–2229. MR3984254 https://doi.org/10.1214/18-AAP1449

[20] JONASSON, J. (2012). Mixing times for the interchange process. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** 667–683. MR3069380

[21] KIPNIS, C., OLLA, S. and VARADHAN, S. R. S. (1989). Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.* **42** 115–137. MR0978701 https://doi.org/10.1002/cpa.3160420202

[22] KOZMA, G. (2007). On the precision of the spectral profile. *ALEA Lat. Am. J. Probab. Math. Stat.* **3** 321–329. MR2372888

[23] LABBÉ, C. and LACOIN, H. (2018). Mixing time and cutoff for the weakly asymmetric simple exclusion process. Preprint. Available at arXiv:1805.12213.

[24] LABBÉ, C. and LACOIN, H. (2019). Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *Ann. Probab.* **47** 1541–1586. MR3945753 https://doi.org/10.1214/18-AOP1290

[25] LACOIN, H. (2015). A product chain without cutoff. *Electron. Commun. Probab.* **20** no. 19, 9. MR3320407 https://doi.org/10.1214/ECP.v20-3765

[26] LACOIN, H. (2016). The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.* **44** 3399–3430. MR3551201 https://doi.org/10.1214/15-AOP1053

[27] LACOIN, H. (2016). Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.* **44** 1426–1487. MR3474475 https://doi.org/10.1214/15-AOP1004

[28] LACOIN, H. and LEBLOND, R. (2011). Cutoff phenomenon for the simple exclusion process on the complete graph. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** 285–301. MR2869447

[29] LEE, T.-Y. and YAU, H.-T. (1998). Logarithmic Sobolev inequality for some models of random walks. *Ann. Probab.* **26** 1855–1873. MR1675008 https://doi.org/10.1214/aop/1022855885

[30] LEVIN, D. A. and PERES, Y. (2016). Mixing of the exclusion process with small bias. *J. Stat. Phys.* **165** 1036–1050. MR3575636 https://doi.org/10.1007/s10955-016-1664-z

[31] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*. 2nd ed. Amer. Math. Soc., Providence, RI. MR3726904

[32] 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 https://doi.org/10.1007/978-3-662-03990-8

[33] LIGGETT, T. M. (2005). *Interacting Particle Systems. Classics in Mathematics*. Springer, Berlin. Reprint of the 1985 original. MR2108619 https://doi.org/10.1007/b138374

[34] LYONS, R., MORRIS, B. J. and SCHRAMM, O. (2008). Ends in uniform spanning forests. *Electron. J. Probab.* **13** 1702–1725. MR2448128 https://doi.org/10.1214/EJP.v13-566

[35] LYONS, R. and OVEIS GHARAN, S. (2018). Sharp bounds on random walk eigenvalues via spectral embedding. *Int. Math. Res. Not. IMRN* **24** 7555–7605. MR3892273 https://doi.org/10.1093/imrn/rnx082

[36] LYONS, R. and PERES, Y. (2016). *Probability on Trees and Networks. Cambridge Series in Statistical and Probabilistic Mathematics* **42**. Cambridge Univ. Press, New York. MR3616205 https://doi.org/10.1017/9781316672815

[37] MORRIS, B. (2006). The mixing time for simple exclusion. *Ann. Appl. Probab.* **16** 615–635. MR2244427 https://doi.org/10.1214/105051605000000728

[38] MORRIS, B. and PERES, Y. (2005). Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields* **133** 245–266. MR2198701 https://doi.org/10.1007/s00440-005-0434-7

[39] OLIVEIRA, R. I. (2013). Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk. *Ann. Probab.* **41** 871–913. MR3077529 https://doi.org/10.1214/11-AOP714

[40] PETE, G. (2008). A note on percolation on $\mathbb{Z}^d$: Isoperimetric profile via exponential cluster repulsion. *Electron. Commun. Probab.* **13** 377–392. MR2415145 https://doi.org/10.1214/ECP.v13-1390

[41] QUASTEL, J. (1992). Diffusion of color in the simple exclusion process. *Comm. Pure Appl. Math.* **45** 623–679. MR1162368 https://doi.org/10.1002/cpa.3160450602

[42] WILSON, D. B. (2004). Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.* **14** 274–325. MR2023023 https://doi.org/10.1214/aoap/1075828054

[43] YAU, H.-T. (1997). Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probab. Theory Related Fields* **109** 507–538. MR1483598 https://doi.org/10.1007/s004400050140