The Annals of Probability
2017, Vol. 45, No. 2, 1003–1070
DOI: 10.1214/15-AOP1080

MIXING TIMES FOR A CONSTRAINED ISING PROCESS ON THE TORUS AT LOW DENSITY

BY NATESH S. PILLAI[1] AND AARON SMITH[2]

Harvard University and University of Ottawa

We study a kinetically constrained Ising process (KCIP) associated with a graph G and density parameter p; this process is an interacting particle system with state space $\{0,1\}^G$, the location of the particles. The number of particles at stationarity follows the Binomial$(|G|, p)$ distribution, conditioned on having at least one particle. The "constraint" in the name of the process refers to the rule that a vertex cannot change its state unless it has at least one neighbour in state "1". The KCIP has been proposed by statistical physicists as a model for the glass transition, and more recently as a simple algorithm for data storage in computer networks. In this note, we study the mixing time of this process on the torus $G = \mathbb{Z}_L^d$, $d \geq 3$, in the low-density regime $p = \frac{c}{|G|}$ for arbitrary $0 < c < \infty$; this regime is the subject of a conjecture of Aldous and is natural in the context of computer networks. Our results provide a counterexample to Aldous' conjecture, suggest a natural modification of the conjecture, and show that this modification is correct up to logarithmic factors. The methods developed in this paper also provide a strategy for tackling Aldous' conjecture for other graphs.

CONTENTS

Received January 2015; revised November 2015.
[1]Supported by an ONR grant.
[2]Supported by a grant from NSERC.
MSC2010 subject classifications. Primary 60J10; secondary 60J20.
Key words and phrases. Ising models, interacting particle systems, mixing times, coalescent process, north-east model.

1. Introduction. The kinetically constrained Ising process (KCIP) refers to a class of interacting particle systems introduced by physicists in [13, 14] to study the glass transition. Versions of this process have accrued other names since then, including the kinetically constrained spin model, the east model [2] and the north-east model [31]. These models have attracted a great deal of interest recently, including applications to combinatorics, computer science and other areas; [6, 7] have useful surveys of places that the KCIP has appeared outside of the physics literature. Recent mathematical progress has included new bounds on the mixing properties of the KCIP in various regimes [3–7, 21, 24]. For a more complete review of recent progress on KCIP within the physics community, see the survey [16] and the references therein.

In this note, we study a simple discrete-time version of this process, though our main result applies, after suitable time scaling, to standard continuous-time analogues as well. Fix a graph $G = (V, E)$ and a density parameter $0 < p < 1$. For a set S, we denote by $\mathrm{Unif}(S)$ the uniform distribution on S. Define a reversible Markov chain $\{X_t\}_{t\in\mathbb{N}}$ on the set of $\{0, 1\}$-labellings of G as follows. To update the chain X_t, choose

$$\begin{aligned} v_t &\sim \mathrm{Unif}(V), \\ p_t &\sim \mathrm{Unif}\big([0, 1]\big). \end{aligned} \tag{1.1}$$

If there exists $u \in V$ such that $(u, v_t) \in E$ and $X_t[u] = 1$, set $X_{t+1}[v_t] = 1$ if $p_t \le p$ and set $X_{t+1}[v_t] = 0$ if $p_t > p$. If no such $u \in V$ exists, set $X_{t+1}[v_t] = X_t[v_t]$. In either case, set $X_{t+1}[w] = X_t[w]$ for all $w \in V \setminus \{v_t\}$.

The state space for the KCIP $\{X_t\}_{t\in\mathbb{N}}$ on a graph G is $\Omega = \{0, 1\}^G$. Set $|V| = n$; for general points $x \in \{0, 1, 2, \ldots\}^G$, write $|x| = \sum_{v\in G} \mathbf{1}_{x[v]\neq 0}$. Let π denote the stationary distribution of $\{X_t\}_{t\in\mathbb{N}}$. For $y \in \Omega$, this is given by

$$\pi(y) = \frac{1}{\mathcal{Z}_{\mathrm{KCIP}}} p^{|y|}(1 - p)^{n-|y|}\mathbf{1}_{|y|>0}, \tag{1.2}$$

where $\mathcal{Z}_{\mathrm{KCIP}} = 1 - (1 - p)^n$ is the normalising constant [see formulas (4.1) and (4.2) below]. Thus, $\pi(y)$ is proportional to the Binomial(n, p) distribution on the

number of nonzero labels in $y \in \Omega$, conditional on having at least one nonzero entry.

In this paper, we study a conjecture stated by David Aldous in [1] about the mixing time of this process. To state Aldous' conjecture, we recall some standard notation that will be used throughout the paper. For sequences $x = x(n)$, $y = y(n)$ indexed by $\mathbb{N}$, we write $y = O(x)$ for $\sup_n \frac{|y(n)|}{|x(n)|} \le C < \infty$ and $y = o(x)$ for $\limsup_{n\to\infty} \frac{|y(n)|}{|x(n)|} = 0$.

We denote by $\mathcal{L}(X)$ the distribution of a random variable X. Recall that for distributions μ, ν on a common measure space $(\Theta, \mathcal{A})$, the *total variation* distance between μ and ν is given by

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup_{A\in\mathcal{A}} \big(\mu(A) - \nu(A)\big).$$

The *mixing profile* for the KCIP Markov chain $\{X_t\}_{t\in\mathbb{N}}$ on Ω with stationary distribution π is given by

$$\tau(\varepsilon) = \inf\Big\{t > 0 : \sup_{X_0=x\in\Omega} \big\|\mathcal{L}(X_t) - \pi\big\|_{\mathrm{TV}} < \varepsilon\Big\}$$

for $0 < \varepsilon < 1$. As usual, the *mixing time* is defined as $\tau_{\mathrm{mix}} = \tau(\frac{1}{4})$. Aldous' conjecture is [1].

CONJECTURE 1 (Aldous). *The mixing time τ_{mix} of the constrained Ising process with parameter p on graph G is $O(p^{-1}|E|\tau_{\mathrm{mix}}^{\mathrm{RW}})$, where $\tau_{\mathrm{mix}}^{\mathrm{RW}}$ is the mixing time of the $\frac{1}{2}$-lazy simple random walk on the graph G.*

Although Conjecture 1 is quite general, it was made in the context of studying the KCIP on a sequence of graphs $\{G_n\}_{n\in\mathbb{N}}$ with associated density $p = p_n = \frac{c}{|G_n|}$ for some fixed $0 < c < \infty$ [1]. This scaling regime for p_n is natural for studying the low-temperature limit of the physical process and has been referred to as the natural equilibrium scale [23]; however, its motivation in [1] is as a model for data storage in computer networks rather than as a model for physical processes. We point out that, although this conjecture seems to be supported in the high-temperature (i.e., $p_n = p$ is constant) regime, substantial evidence (both theoretical and numerical) has been collected since 2002 that the conjecture is not correct in the low-temperature regime that we study. For instance, the qualitative results of [19] suggest that this conjecture is not correct, though the authors in [19] do not give a proof of this statement.

For a positive integer $L \in \mathbb{N}$, let $\Lambda(L, d)$ denote the d-dimensional torus with $n = L^d$ points; this is a Cayley graph with vertex set, generating set and edge set given by

$$\begin{aligned} V &= \mathbb{Z}_L^d, \\ \mathrm{Gen} &= \big\{(1,0,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,0,0,\ldots,1)\big\}, \\ E &= \big\{(u,v) \in V \times V : u - v \in \pm\mathrm{Gen}\big\}. \end{aligned}$$

Set

$$n = \big|\Lambda(L,d)\big| = L^d.$$

In this paper, we study the KCIP on a sequence of graphs $\{\Lambda(L,d)\}_{L\in\mathbb{N}}$ with density

$$p = p_n = \frac{c}{n} \tag{1.3}$$

for some fixed constant $0 < c < \infty$ and fixed dimension $d \geq 3$; see Section 9 for a brief explanation of how our method applies when d and c are allowed to vary with L.

The mixing time of the simple random walk on $G = \Lambda(L,d)$ is known to be $\tau_{\text{mix}}^{\text{RW}} \approx n^{\frac{2}{d}}$ (see, e.g., Theorem 5.5 of [22]). Thus, Aldous' conjecture for $G = \Lambda(L,d)$ suggests a mixing time of $\tau_{\text{mix}} = O(n^{2+\frac{2}{d}})$. This is correct for $d = 1$, but we show in Theorem 2 below that this conjecture is incorrect for $d \geq 3$.

The following is our main result, in which we prove a modified version of Conjecture 1 for the torus.

THEOREM 2 (Mixing of the constrained Ising process on the torus). *Fix* $0 < c < \infty$ *and* $d \geq 3$. *For* $p = p_n$ *defined in* (1.3), *the mixing time of the* KCIP *on* $\Lambda(L,d)$ *satisfies*

$$C_1 n^3 \leq \tau_{\text{mix}} \leq C_2 n^3 \log(n)$$

for some constants C_1, C_2 *that may depend on* c, d *but are independent of* n.

REMARK 1.1. We conjecture that the lower bound is tight, that is, $\tau_{\text{mix}} \approx n^3$ for $d \geq 3$. We also conjecture that $\tau_{\text{mix}} \approx n^3 \log(n)$ for $d = 2$. These different predictions come from the same source: in all dimensions, the main obstacle to mixing in the KCIP can be viewed as the time it takes for "many" particles to "collide" with each other. We make this intuition formal in Section 6 by relating the KCIP to coalescing random walks. This formalization suggests that the number of steps required for "many collisions" to occur in the KCIP should be very similar to the consensus time for the voter model. After appropriate rescaling, the main results in [9] say that this consensus time scales like n^3 for $d \geq 3$ but at the slightly faster rate of $n^3 \log(n)$ for $d = 2$.

In the statement of Theorem 2 and throughout the paper, we assume that both the quantities $0 < c < \infty$ and $3 \leq d \in \mathbb{N}$ are fixed; only n grows. In particular, in Theorem 2 and all other calculations, bounds that are "uniform" are implied to be uniform only in n and other explicitly mentioned variables; they will generally not be uniform in c or d. Throughout the paper, we will denote by C a generic constant, whose value may change from one occurrence to the next, but is independent of n.

1.1. *Relationship to previous work.* Although the KCIP we study was introduced in the physics literature [13, 14] and discussed in later work such as [1], most recent mathematical work on mixing bounds for KCIPs has focused on different local constraints, which give rise to qualitatively different behaviour. Thus, even if we studied the same regime, our results would not imply (or be implied by) recent work in this area. However, our primary contribution to the literature is the fact that our work is in a new regime: we obtain good mixing bounds on a KCIP that apply in the regime of high dimension $d \geq 3$, low density $p_n \approx n^{-1}$, and under the strong metric $\|\cdot\|_{\mathrm{TV}}$. All three of these distinctions can make the problem harder than working in dimension $d = 1$, at high density $p_n = p$, or in a weaker metric.

We briefly review some recent work on the mixing properties of related constrained Ising processes [3–7, 21, 24], contrasting our work with specific papers. Many previous results, such as [5, 7], deal primarily with the regime in which p is a constant, independently of n. In this regime, many KCIPs mix relatively quickly and the obstacles to mixing are quite different. Other results, such as [2, 15], study the small-p regime, but only in one dimension. In particular, the methods employed in [2] completely break down for $d > 1$ and thus are not applicable to our setting. The recent paper [6] seems most similar to ours. In [6], the authors study the mixing of a related KCIP on $\Lambda(L, d)$ at density $p \approx \frac{1}{n}$ and obtain results in greater generality than ours; these results are also qualitatively similar. However, the authors of [6] focus on bounding the relaxation time of the process, rather than the mixing time; the bounds obtained in [6] cannot be used to obtain sharp estimates on the mixing time. We call additional attention to [19]. Like our paper and [6], the authors in [19] study the properties of the KCIP at density $p \approx \frac{1}{n}$. Furthermore, as in our paper, a key step of the analysis in [19] is the construction of a mapping between the KCIP and a coalescent process. However, the details of their proofs are quite different (they use a different coalescent process, and construct an exact mapping between generators rather than an approximate coupling of processes), as are their final results.

1.2. *Outline for the paper.* This paper is largely devoted to the proof of Theorem 2. In Section 2, we set up notation and give a proof sketch. In Section 3, we give a general upper bound on the mixing properties of Markov chains on finite state spaces; the rest of the proof of the upper bound in Theorem 2 consists of estimating the constants in this general bound. In Section 4, we prove the lower bound in Theorem 2. Sections 5 to 7 contain most of the work required to prove the upper bound in Theorem 2. In Section 5, we study the behaviour of the KCIP at low density by comparing it to the simple exclusion process. In Sections 6 and 7, we detail the behaviour of the KCIP at high density by comparing it to the coalescent process. In Section 8, we combine the results obtained in earlier sections and give the proof of Theorem 2. Finally, in Section 9 we discuss related problems and the extent to which our methods apply to them.

2. A roadmap for the proof. We first explain the heuristic arguments that make Conjecture 1 plausible and point out the key time scales involved. For the reader's convenience, we also give a proof sketch highlighting all the major steps involved in the proof of Theorem 2.

2.1. *Heuristics and key time scales.* We discuss heuristics for the time scales on which important changes to the KCIP occur.

1. If the initial state X_0 has a large number of particles, the number of particles remaining will generally be cut in half every $O(n^3)$ steps. This observation is crucial to our proof. We show this bound using closely related bounds on the "coalescence time" of a collection of random walkers on a graph (see [9]). This bound implies that, after an initial transient period of at most $O(n^3 \log(n))$ steps, the number of particles in the KCIP will be $O(1)$. This is the only place we obtain an extra $\log n$ factor in our proof of the upper bound of Theorem 2. We suspect that, in fact, this transient period is only of total length $O(n^3)$.
2. After the initial transient period, the KCIP generally has $O(1)$ well-separated particles. Every $O(n^2)$ steps, a particle will spawn a neighbour; one of these neighbouring particles will be removed in $O(n)$ steps. Ignoring the times at which any particle has a neighbour, the $O(1)$ well-separated particles will appear to be evolving according to an independent random walk on $\Lambda(L, d)$, slowed down by a factor of roughly n^2; see Figure 1. Since all particles have no neighbours at most times under the stationary measure, ignoring the times at which any particle has a neighbour does not greatly influence one's view of the process.
3. After the initial transient period, and again ignoring the times at which any particle has a neighbour, it often takes $O(n^3)$ steps to decrease the number of particles by 1. To see this, note that the number of particles can only decrease when two existing particles "collide." Recall that the expected collision time for two random walkers on the torus is $O(n)$. By the above heuristic, the particles in the KCIP are undergoing simple random walk that is slowed down by a factor of n^2; thus, the expected time to a collision is $O(n^3)$. It also turns out that collisions "often" result in the number of particles being decreased.

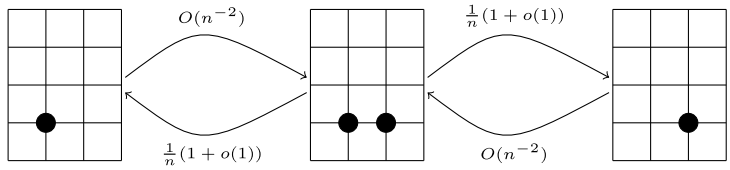

FIG. 1. *Simple Random Walk Heuristic*: *The middle configuration lasts for a very short time compared to the outer configurations*; *ignoring the middle "transient" configuration*, *the KCIP particle seems to be undergoing simple random walk.*

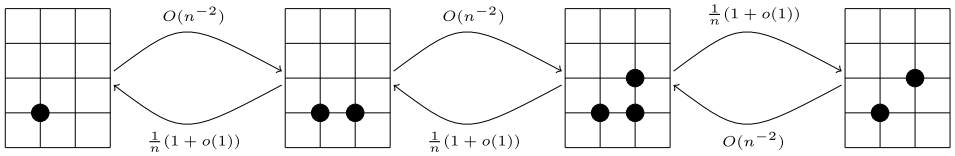

FIG. 2. *Particle Creation Heuristic*: *The most common sequence of configurations leading from a single particle to a pair of nonadjacent particles.*

4. After the initial transient period, and again ignoring the times at which any particle has a neighbour, it takes $O(n^3)$ steps to increase the number of particles by 1. Indeed, increasing the number of nonadjacent particles by 1 requires an intermediate time at which three particles are "touching" in $\Lambda(L,d)$, where two of them share the same neighbour. It takes $O(n^2)$ steps for a particle to spawn a neighbouring particle, and whenever two particles are adjacent, the probability of one of these two particles being removed before a third "touching" particle is added is $1-O(n^{-1})$. Thus, it takes $O(n^3)$ steps to obtain three "touching" particles; see Figure 2 for a generic illustration of how this happens. It is easy to check that a configuration with three "touching" particles often degenerates into one with two nonadjacent particles after $O(n)$ steps.

The basis of Conjecture 1 is heuristic (2) above: individual particles at distance greater than one in the KCIP on $\Lambda(L,d)$ tend to behave like independent random walkers on $\Lambda(L,d)$, slowed down by a factor of roughly n^2. Thus, for density $p=\frac{c}{n}$, we might expect the KCIP to have roughly c particles during most times and to behave quite similarly to the simple exclusion process (SEP) with c particles (see [8, 17] for an introduction). Heuristics (1) and (4) explain why Conjecture 1 is not telling the whole story: heuristic (1) points out that it takes a long time to go from n particles to $O(1)$ particles, while heuristic (4) points out that it takes a long time to go from 1 particle to 2 well-separated particles. The lower bound in Theorem 2 is obtained by making heuristic (4) rigorous.

An obvious modification to Conjecture 1 is that the mixing time of the KCIP is at most the maximum of these three time scales, and this is the approach we take in this paper. There are essentially three obstacles to making this approach rigorous. The first is to deal with the fact that we would like to compare a single KCIP to many different SEPs—the SEP with c particles, but also the SEP with $c-1$ particles, $c+1$ particles, *etc.* Breaking apart the KCIP in this way is the subject of Step 1 and Step 4 in the proof sketch below. The second is to complete the comparison of the SEP to a suitably tamed version of the KCIP. This is the subject of Step 2 below. The third is to ensure that a KCIP started with a large number ($\gg c$) particles quickly enters a state with roughly c particles. This corresponds to Step 3 below.

2.2. *Proof sketch.* Our methods in some of the steps in the following outline are applicable for a generic graph G, and the others are specific to the torus $\Lambda(L,d)$. For $1 \le k \le \frac{n}{2}$, let $\Omega_k \subset \Omega$ be configurations of k particles for which no two particles are adjacent, that is,

$$\Omega_k = \Big\{ X \in \{0,1\}^G : \sum_{v \in V} X[v] = k, \sum_{(u,v)\in E} X[u]X[v] = 0 \Big\}. \tag{2.1}$$

Also set $\Omega' = \Omega \setminus \bigcup_{k=1}^{\frac{n}{2}} \Omega_k$. For each $k \le \frac{n}{2}$, we will denote by $\tau_{\text{mix}}^{(k)}$ the mixing time of X_t "restricted to" Ω_k (our notion of "restriction" is defined more carefully in Section 5). Define the quantity

$$\text{Occ}_k(\varepsilon, N) = \sup_{x\in\Omega} \inf \Big\{ T \ge 1 : X_0 = x, \mathbb{P}\Big(\sum_{s=0}^{T} 1_{X_s \in \Omega_k} > N \Big) > 1 - \varepsilon \Big\}.$$

For a fixed N and small ε, $\text{Occ}_k(\varepsilon, N)$ denotes the first (random) time at which the occupation measure of X_t in Ω_k exceeds N with high $(1-\varepsilon)$ probability. For $x \in \Omega_k$, define the exit time:

$$L_k(x) = \inf \Big\{ t : X_0 = x, X_t \in \bigcup_{j \ne k} \Omega_j \Big\}. \tag{2.2}$$

Our proof strategy for the upper bound in Theorem 2 entails the following steps.

Step 1. We show that for a universal constant $k_{\max} = k_{\max}(G, c)$ depending only on the local structure of the graph G and the constant c from (1.3),

$$\tau_{\text{mix}} = O\Big(\sup_{1\le k\le k_{\max}} \tau_{\text{mix}}^{(k)} + \sup_{1\le k\le k_{\max}} \text{Occ}_k\Big(\frac{1}{8k_{\max}}, C\tau_{\text{mix}}^{(k)} \Big) \Big).$$

This is an immediate consequence of Lemma 3.1.

Step 2. By a comparison argument using the simple exclusion process, we show that $\tau_{\text{mix}}^{(k)} = O((\alpha_{n,k}^{\text{SE}})^{-1} n \log n)$ uniformly in $1 \le k \le k_{\max}$, where $\alpha_{n,k}^{\text{SE}}$ is the log-Sobolev constant of the simple exclusion process on G with k particles. For $G = \Lambda(L,d)$, it is known that $(\alpha_{n,k}^{\text{SE}})^{-1} = O(n^{1+\frac{2}{d}})$ (see [32]), and thus $\tau_{\text{mix}}^{(k)} = O(n^{2+\frac{2}{d}} \log n)$ uniformly in $1 \le k \le k_{\max}$ for the KCIP on $\Lambda(L,d)$. See Lemma 5.1.

Step 3. For $G = \Lambda(L,d)$, by coupling the KCIP to a "colored" version of the coalescence process over short time periods, we show that the process

$$V_t = \sum_{v\in V} X_t[v] \tag{2.3}$$

satisfies the "drift condition"

$$\mathbb{E}[V_{t+\varepsilon n^3} - V_t | X_t] \le -\delta V_t + C$$

for some $C < \infty$ and $\varepsilon, \delta > 0$, all independent of n. See Theorem 6.1.

Step 4. For $G = \Lambda(L, d)$, by direct calculation and further comparison to a coalescence process, we show that

$$\inf_{x \in \Omega_k} \mathbb{P}\big[L_k(x) > C_1 n^3\big] > \varepsilon > 0$$

and

$$\sup_{x \in \Omega_k} \mathbb{P}\big[L_k(x) < C_2 n^3\big] > \varepsilon > 0$$

for some constants $C_1, C_2, \varepsilon > 0$ independent of n. See Lemmas 7.3 and 7.6.

Step 5. Conclude from Step 3 and Step 4 that $\sup_{1 \le k \le k_{\max}} \mathrm{Occ}_k(\frac{1}{8k_{\max}}, C\tau_{\mathrm{mix}}^{(k)}) = O(n^3 \log(n))$. See Lemma 8.1.

The lower bound is obtained in Theorem 2 by a direct computation of the expected time for the KCIP $\{X_t\}_{t \ge 0}$ on $\Lambda(L, d)$ to first have two nonadjacent particles when started from a single particle. Some of the steps above that are specific to the torus $\Lambda(L, d)$ can be extended for other graphs as well, as discussed in Section 9.

3. General mixing bounds for decomposable Markov chains. In this section, we give a general bound on the mixing time of decomposable Markov chains. This result will be later applied to the KCIP Markov chain to achieve Step 1 in the proof of Theorem 2. The bounds in this section apply to Markov chains other than the KCIP, and thus are of independent interest; they are developed further and applied to other examples in a companion paper [29].

Consider an aperiodic, irreducible and reversible Markov chain $\{Z_t\}_{t \in \mathbb{N}}$ with transition matrix K and stationary distribution $\tilde{\pi}$ on a finite state space Θ. Our goal is to bound the mixing time of $\{Z_t\}_{t \in \mathbb{N}}$ in terms of the mixing times of various *restricted* and *projected* chains; these L^1 mixing bounds are loosely analogous to the L^2 bounds in [20]. We begin by fixing n and writing the partition $\Theta = \bigsqcup_{k=1}^{n} \Theta_k$. For $1 \le k \le n$, set $\eta_k(0) = 0$ and for $s \in \mathbb{N}$, recursively define the sequences of times

$$\begin{aligned} \eta_k(s) &= \inf\big\{t > \eta_k(s-1) : Z_t \in \Theta_k\big\}, \\ \kappa_k(s) &= \sup\big\{u : \eta_k(u) \le s\big\}. \end{aligned} \tag{3.1}$$

The quantity κ_k can also be written as

$$\kappa_k(T) = \sum_{t=1}^{T} \mathbf{1}_{Z_t \in \Theta_k}. \tag{3.2}$$

Both η, κ depend on the initial condition Z_1. We also define the associated *restricted* processes $\{Z_t^{(k)}\}_{t \in \mathbb{N}}$ by

$$Z_t^{(k)} = Z_{\eta_k(t)}. \tag{3.3}$$

This process is called the *trace* of $\{Z_t\}_{t\in\mathbb{N}}$ on Θ_k. Since $\{Z_t\}_{t\in\mathbb{N}}$ is recurrent, we have for all $t \in \mathbb{N}$ that $\eta_k(t) < \infty$ almost surely, and so $Z_t^{(k)}$ is almost surely well-defined for all $t \in \mathbb{N}$. The process $\{Z_t^{(k)}\}_{t\in\mathbb{N}}$ is a Markov chain on Θ_k; denote by K_k the associated transition kernel on Θ_k. The kernel K_k inherits aperiodicity, irreducibility and reversibility from K and its stationary distribution is given by $\tilde{\pi}_k(A) = \frac{\tilde{\pi}(A)}{\tilde{\pi}(\Theta_k)}$ for all $A \subset \Theta_k$. We denote by φ_k the mixing time of the kernel K_k.

Say that the chain $\{Z_t\}_{t\geq 0}$ is *sufficiently lazy* if either of the following two conditions hold:

1. $\min_{x\in\Theta} K[x, x] \geq \frac{1}{2}$, or
2. there exists a set $S \subset \Theta$ and $\delta < \frac{1}{24}$ satisfying:

$$\pi(S) \geq 1 - \delta, \qquad \min_{x\in S} K[x, x] \geq 1 - \delta,$$

$$\max_{x\notin S} \sum_{y\in S} K[x, y] \leq \delta.$$

For $0 < a < \frac{1}{2}$, define the universal constants c_a and c'_a as in Theorem 1.1 of [28]. The following simple bound on the mixing time of K combines Lemmas 3.1 and A.2 of [29].

LEMMA 3.1 (Basic mixing bound). *Fix* $0 < a < \frac{1}{2}$ *and* $1 - a < \beta < 1$. *Define* $\gamma = \min(\frac{1}{2}, \frac{a+\beta-1}{\beta}) > 0$ *and fix a collection of indices* $I \subset \{1, 2, \ldots, n\}$ *satisfying*

$$\sum_{k\in I} \tilde{\pi}(\Theta_k) > \beta.$$

Set

$$\mathcal{T} \equiv \inf\left\{T : \inf_{0<t<T} \sup_{k\in I}\left(\frac{c'_\gamma \varphi_k}{t} + \sup_{z\in\Theta} \mathbb{P}_z[\kappa_k(T) < t]\right) < \frac{1}{4}\right\}.$$

Then the mixing time τ^Z_{mix} *of any sufficiently lazy chain* $\{Z_t\}_{t\in\mathbb{N}}$ *satisfies*

$$\frac{\tau^Z_{\text{mix}}}{c_a} \leq \frac{4}{3} c_a \mathcal{T}.$$

Thus, Lemma 3.1 relates the mixing time of a Markov chain to the mixing times of its traces on subsets of the state space with reasonably large stationary measure and the corresponding occupation times on those subsets. This is possible due to the remarkable results obtained in [25, 28], where the authors obtain a bound on the mixing times of *reversible* Markov chains in terms of hitting times.

REMARK 3.2. Since our mixing bounds in Section 5 below are obtained by a bound on the log-Sobolev constant of various restricted chains, and our bound

on occupation measure can easily be translated into a bound on the spectral gap of an associated projected chain, the reader may ask why we do not use the well developed theory in [20] for decomposable Markov chains to bound the mixing of a Markov chain in terms of restricted and projected chains. One reason is convenience: unlike the bounds in [20], Lemma 3.1 requires only a bound on φ_k for *some* k, not *all* k. Since the bounds in Section 5 apply only for k small, Lemma 3.1 allows us to avoid doing the substantial extra work of finding bounds on φ_k for k large. The second, and more important, reason is that it is impossible to get an upper bound on the mixing time that is smaller than $O(n^4 \log(n))$ using bounds from [20] and any partition of Ω similar to the partition that we use. The extra factor of n comes primarily from the fact that the probability of moving from Ω_{k+1} to Ω_k within $O(n^2)$ steps is very far from uniform over starting points $x \in \Omega_{k+1}$. See [29] for more examples illustrating this point.

4. Lower bound for the mixing time of KCIP on $\Lambda(L, d)$. In this section, we give a direct computation that leads to a lower bound on the mixing time of the KCIP on any m-regular graph $G = (V, E)$ that contains no triangles. The torus $\Lambda(L, d)$ is $2d$-regular and triangle-free, and this bound will immediately give the lower bound in Theorem 2. Before proceeding, we verify that the stationary distribution of the KCIP is given by formula (1.2). If $\mathbb{P}[X_{t+1} = y | X_t = x] > 0$, then either

$$\begin{aligned} \mathbb{P}[X_{t+1} = y | X_t = x] &= \frac{p}{1-p} \mathbb{P}[X_{t+1} = x | X_t = y] \\ &= \frac{\pi(y)}{\pi(x)} \mathbb{P}[X_{t+1} = x | X_t = y] \end{aligned} \tag{4.1}$$

or

$$\begin{aligned} \mathbb{P}[X_{t+1} = y | X_t = x] &= \frac{1-p}{p} \mathbb{P}[X_{t+1} = x | X_t = x] \\ &= \frac{\pi(y)}{\pi(x)} \mathbb{P}[X_{t+1} = x | X_t = y], \end{aligned} \tag{4.2}$$

depending on whether $\sum_{v \in G} x[v] < \sum_{v \in G} y[v]$ or not. Thus, the Markov chain $\{X_t\}_{t \in \mathbb{N}}$ satisfies the detailed balance equation with respect to π, and thus has π as its stationary distribution.

We continue by setting notation that will be used throughout the remainder of the paper. For $u, v \in \Lambda(L, d)$, denote by $|u - v|$ the smallest number of edges needed to traverse from u to v via a connected path; this is the usual graph distance. Let

$$\mathcal{B}_\ell(v) = \{w \in \Lambda(L, d) : |v - w| \leq \ell\} \tag{4.3}$$

be the ball of radius ℓ around v in the graph distance. For $t \in \mathbb{N}$, let G_t be the subgraph of G induced by the vertices $\{v \in G : X_t[v] = 1\}$, with vertices $V(G_t)$

and edges $E(G_t)$. For a vertex $u \in \Lambda(L,d)$, define $\mathrm{Comp}_t(u)$ to be the collection of vertices contained in the same connected component as u in G_t, define Comp_t to be the collection of distinct connected components in G_t, and let $Y_t = |\mathrm{Comp}_t|$ be the number of connected components in G_t. Recall that $V_t = \sum_{v\in G} X_t[v] = |G_t|$ is the number of vertices having state 1. For a KCIP started at time 0 with $V_0 = 1$, define the associated *triple* time by

$$\zeta_{\text{triple}} = \inf\{t : V_t \geq 3\}. \tag{4.4}$$

LEMMA 4.1 (Component growth). *Let G be an m-regular graph ($m > 1$) with no triangles. Fix $\varepsilon > 0$ and assume that $V_0 = 1$. Then*

$$\mathbb{P}\left[\zeta_{\text{triple}} < \varepsilon \frac{n^3}{3c^2 m(m-1)}\right] = O(\varepsilon), \tag{4.5}$$

where the implied constant is uniform over $0 < \varepsilon < \varepsilon_0$ sufficiently small and does not depend on G or c.

PROOF. Define the matrix

$$K = \begin{pmatrix} 1 - \dfrac{cm}{n^2} & \dfrac{cm}{n^2} & 0 \\ \dfrac{2}{n}\left(1 - \dfrac{c}{n}\right) & 1 - \dfrac{c(2m-2)}{n^2} - \dfrac{2}{n}\left(1 - \dfrac{c}{n}\right) & \dfrac{c(2m-2)}{n^2} \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.6}$$

It is straightforward to check that, for $0 \leq s < \zeta_{\text{triple}}$ and $a \in \{1,2,3\}$,

$$\mathbb{P}[V_{s+1} = a | X_s] = K[V_s, a]. \tag{4.7}$$

Most significantly, X_s appears on the right-hand side only through V_s. We justify this by considering the various cases. If $V_s = 1$, the transition probabilities $\mathbb{P}[V_{s+1} = i | X_s]$ depend only on the number of vertices labelled 1 at time s (this is 1) and the number of vertices adjacent to this vertex (since our graph is m-regular, this is m). If $V_s = 2$, the transition probabilities $\mathbb{P}[V_{s+1} = i | X_s]$ depend only on the number of vertices labelled 1 at time s (this is 2), the number of vertices labelled 1 adjacent to other vertices labelled 1 (this is also 2) and the number of vertices labelled 0 adjacent to vertices labelled 1 (since our graph is both m-regular and triangle-free, this is $2m-2$).

By direct computation,

$$\mathbb{E}_1[\zeta_{\text{triple}}] = \frac{n^2}{cm} + \mathbb{E}_2[\zeta_{\text{triple}}],$$

$$\left(\frac{2}{n}\left(1 - \frac{c}{n}\right) + \frac{c(2m-2)}{n^2}\right)\mathbb{E}_2[\zeta_{\text{triple}}] = 1 + \frac{2}{n}\left(1 - \frac{c}{n}\right)\mathbb{E}_1[\zeta_{\text{triple}}],$$

and so we obtain

$$\mathbb{E}_1[\zeta_{\text{triple}}] = \frac{n^3}{c^2 m(m-1)} + O(n^2). \tag{4.8}$$

Next, observe that for any $T \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}[\zeta_{\text{triple}} > 2T] &= \mathbb{P}[\zeta_{\text{triple}} > 2T \,|\, \zeta_{\text{triple}} > T]\mathbb{P}[\zeta_{\text{triple}} > T] \\ &= \mathbb{E}\big[\mathbb{P}[\zeta_{\text{triple}} > 2T \,|\, \zeta_{\text{triple}} > T] \,|\, X_T, \zeta_{\text{triple}} > T\big]\mathbb{P}[\zeta_{\text{triple}} > T] \\ &\le \mathbb{P}[\zeta_{\text{triple}} > T]^2.\end{aligned}$$

By the same calculation, we have for any $k, T \in \mathbb{N}$ that

$$\mathbb{P}[\zeta_{\text{triple}} > kT] \le \mathbb{P}[\zeta_{\text{triple}} > T]^k.$$

It is also clear that $\mathbb{P}[\zeta_{\text{triple}} > t]$ is a monotonely decreasing function of t.

We now prove our result by contradiction. If we assume that $\mathbb{P}[\zeta_{\text{triple}} < \varepsilon \frac{n^3}{3c^2 m(m-1)}] > 10\varepsilon$ for some $n \in \mathbb{N}$, we would have

$$\begin{aligned}\mathbb{E}[\zeta_{\text{triple}}] &\le \sum_{k=0}^{\infty} \mathbb{P}\left[\zeta_{\text{triple}} > k\left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil\right]\left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil \\ &\le \left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil \sum_{k=0}^{\infty} \mathbb{P}\left[\zeta_{\text{triple}} > \left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil\right]^k \\ &\le \left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil \sum_{k=0}^{\infty} (1-10\varepsilon)^k \\ &= \left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil \frac{1}{10\varepsilon}.\end{aligned}$$

Combining this with inequality (4.8), we have

$$\frac{n^3}{c^2 m(m-1)} + O(n^2) \le \left\lceil \varepsilon \frac{n^3}{3c^2 m(m-1)} \right\rceil \frac{1}{10\varepsilon}.$$

This inequality is clearly false for all sufficiently large n, and so the assumption that $\mathbb{P}[\zeta_{\text{triple}} < \varepsilon \frac{n^3}{3c^2 m(m-1)}] > 10\varepsilon$ must also be false for all sufficiently large n. This completes the proof. □

We now conclude with the lower bound in Theorem 2.

THEOREM 3 (Lower bound on mixing time). *Fix an m-regular ($m > 1$) graph G with no triangles. Then the KCIP on G with success probability $p = \frac{c}{n}$ has mixing time satisfying*

$$\tau_{\text{mix}} \ge C \frac{1}{\mathcal{Z}_c} \frac{n^3}{m(m-1)},$$

where C does not depend on c or G and

$$\mathcal{Z}_c = \frac{3c^2 \max(1, -\log_2((c+\frac{c^2}{2})\frac{e^{-c}}{1-e^{-c}}))}{(c+\frac{c^2}{2})\frac{e^{-c}}{1-e^{-c}}} > 0.$$

PROOF. Fix $\varepsilon > 0$, let $T = \lfloor \varepsilon \frac{n^3}{3c^2 m(m-1)} \rfloor$ and define the set $A = \{x \in \Omega : \sum_{v\in G} x[v] \le 2\}$. Let X_0 be such that $V_0 = 1$. From Lemma 4.1 and equation (1.2) for the stationary distribution π, we calculate:

$$\begin{aligned}
\|\mathcal{L}(X_T) - \pi\|_{\mathrm{TV}} &\ge \pi(A^c) - \mathbb{P}[X_T \in A^c] \\
&\ge \pi(A^c) - \mathbb{P}[\zeta_{\text{triple}} < T] \\
&= 1 - n\frac{c}{n}\left(1-\frac{c}{n}\right)^{n-1}\frac{1}{1-(1-\frac{c}{n})^n} \\
&\quad - \frac{n(n-1)}{2}\frac{c^2}{n^2}\left(1-\frac{c}{n}\right)^{n-2}\frac{1}{1-(1-\frac{c}{n})^n} + O(\varepsilon) \\
&= \left(c+\frac{c^2}{2}\right)\frac{e^{-c}}{1-e^{-c}}(1+o(1)) + O(\varepsilon).
\end{aligned}$$

Thus, for $\varepsilon \ll \frac{1}{4}(c+\frac{c^2}{2})\frac{e^{-c}}{1-e^{-c}}$ sufficiently small,

$$\|\mathcal{L}(X_T) - \pi\|_{\mathrm{TV}} \ge \frac{1}{2}\left(c+\frac{c^2}{2}\right)\frac{e^{-c}}{1-e^{-c}} > 0$$

uniformly in $n > N(\varepsilon)$ sufficiently large. Since the mixing profile satisfies

$$\tau\left(\frac{1}{2}\left(c+\frac{c^2}{2}\right)\frac{e^{-c}}{1-e^{-c}}\right) \le \max\left(1, -\log_2\left(2\left(c+\frac{c^2}{2}\right)\frac{e^{-c}}{1-e^{-c}}\right)\right)\tau_{\mathrm{mix}},$$

by Lemmas 4.11 and 4.12 of [22], this implies

$$\tau_{\mathrm{mix}} \ge \frac{T}{\max(1, -\log_2(2(c+\frac{c^2}{2})\frac{e^{-c}}{1-e^{-c}}))}$$

for ε sufficiently small. This completes the proof. □

5. Mixing times of the trace of KCIP on Ω_k. For the remainder of the paper, we restrict our attention to the partition of the partition of the state space Ω for the KCIP into the states $\{\Omega_k\}_{k=1}^{\lfloor \frac{n}{2} \rfloor}$ defined in equation (2.1) and the remainder term $\Omega' = \Omega \setminus \bigcup_{k=1}^{\lfloor \frac{n}{2} \rfloor} \Omega_k$. In this section, we obtain bounds on the mixing times of the trace of the KCIP on Ω_k. As these trace walks mix substantially more quickly than the KCIP Markov chain X_t on Ω, we do not need these mixing time bounds to be tight.

Fix $1 \le k \le \frac{n}{2}$, and let $Q_{n,k}$ be the kernel of the trace of $\{X_t\}_{t\in\mathbb{N}}$ on Ω_k [recall that the trace of a Markov chain on a subset of its state space is defined in formula (3.3)]. Denote by $\tau_{n,k}$ the mixing time of $Q_{n,k}$. The key result of this section is the following.

LEMMA 5.1 (Mixing time of restricted walks). *With the notation as above,*

$$\tau_{n,k} \le Cn^{2+\frac{2}{d}} \log(n)$$

for some constant $C = C(c, k, d)$ that does not depend on n.

We proceed by using comparison theory, a tool developed for proving mixing bounds for a Markov chain by comparing its transition rates to a similar and better-understood chain (see, e.g., [10] or [12] for an introduction to this method). We will use the simple exclusion process (SE) on $\Lambda(L, d)$ as the basis of our comparison argument (see [8, 17] for an introduction to the simple exclusion process).

DEFINITION 5.2 (Simple exclusion process on $\Lambda(L, d)$). The simple exclusion process $\{Z_t\}_{t\in\mathbb{N}}$ is a Markov chain on the finite state space

$$\Omega_{n,k}^{\mathrm{SE}} \equiv \left\{ Z \in \{0, 1\}^n : \sum_i Z[i] = k \right\}. \tag{5.1}$$

To update Z_t, choose two adjacent vertices $u_t, v_t \in \Lambda(L, d)$ uniformly at random and set

$$Z_{t+1}[u_t] = Z_t[v_t],$$
$$Z_{t+1}[v_t] = Z_t[u_t]$$

and $Z_{t+1}[w] = Z_t[w]$ for $w \notin \{u_t, v_t\}$.

The approach in this section is to first note that the simple exclusion process with k particles has good mixing properties (we use the results in [32], though others would suffice for our purposes) and then use a comparison argument to show that the mixing properties of the trace of the KCIP on Ω_k cannot be much worse. We use the simple exclusion process because it makes both parts of this argument easy: it has already been carefully analyzed, and it is similar enough to the trace of the KCIP that the comparison argument is short and involves only soft arguments. It is likely that the conclusions we need can be achieved by comparison to other processes on the torus.

5.1. *Comparison of Markov chains using Dirichlet forms.* Before stating the main result of this section carefully, for the reader's convenience, we recall some relevant results for comparing Dirichlet forms. We use the bounds in [30], rather

than the similar and simpler results from [10, 12], because we will compare chains (KCIP and SE) with different state spaces; the bounds in [10, 12] cannot be used in this situation.

DEFINITION 5.3 (Norms, forms and related functions). For a general Markov chain on a finite state space X with kernel P and unique stationary distribution π, and any function $f : X \to \mathbb{R}$ that is not identically 0, we respectively define the L_2 norm, variance, Dirichlet form and entropy form as

$$\begin{aligned}
\|f\|_{2,\pi}^2 &= \sum_{x\in X} |f(x)|^2 \pi(x), \\
V_\pi(f) &= \frac{1}{2} \sum_{x,y\in X} |f(x) - f(y)|^2 \pi(x)\pi(y), \\
\mathcal{E}_P(f,f) &= \frac{1}{2} \sum_{x,y\in X} |f(x) - f(y)|^2 P(x,y)\pi(x), \\
L_\pi(f) &= \sum_{x\in X} |f(x)|^2 \log\Big(\frac{f(x)^2}{\|f\|_{2,\pi}^2}\Big)\pi(x).
\end{aligned} \tag{5.2}$$

Recall that the *log-Sobolev constant* of a Markov transition matrix P is given by

$$\alpha(P) = \inf_{f\neq 0} \frac{\mathcal{E}_P(f,f)}{L_\pi(f)}. \tag{5.3}$$

DEFINITION 5.4 (Extensions). Let K, Q be the kernels of two $\frac{1}{2}$-lazy, aperiodic, irreducible, reversible Markov chains. Assume that K has stationary measure μ on a state space $\widehat{\Theta}$ while Q has stationary measure ν on a state space $\Theta \subset \widehat{\Theta}$. Denote by f a function on Θ, and call a function $\widehat{f}$ on $\widehat{\Theta}$ an *extension* of f if $\widehat{f}(x) = f(x)$ for all $x \in \Theta$.

Next, fix a family of probability measures $\{\mathbb{P}_x[y]\}_{x\in\widehat{\Theta}}$ on Θ that satisfy $\mathbb{P}_x(\cdot) = \delta_x(\cdot)$ for $x \in \Theta$. We will use only extensions of the form

$$\widehat{f}(x) = \sum_{y\in\Theta} \mathbb{P}_x[y] f(y). \tag{5.4}$$

We call extensions of the form (5.4) *linear extensions*.

Fix a linear extension. For each pair $(x, y) \in \widehat{\Theta}$ with $K(x, y) > 0$, fix a joint probability distribution $\mathbb{P}_{x,y}$ on $\Theta \times \Theta$ satisfying $\sum_a \mathbb{P}_{x,y}[a, b] = \mathbb{P}_y[b]$ for all $b \in \Theta$ and $\sum_b \mathbb{P}_{x,y}[a, b] = \mathbb{P}_x[a]$ for all $a \in \Theta$. This is a coupling of the distributions $\mathbb{P}_x, \mathbb{P}_y$.

DEFINITION 5.5 (Paths, flows). Finally, for each $a, b \in \Theta$ with $\sum_{x,y\in\widehat{\Theta}} \mathbb{P}_{x,y}[a, b] > 0$, we define a *flow* in Θ from a to b. To do so, call a sequence of vertices $\gamma = [a = v_{0,a,b}, v_{1,a,b}, \ldots, v_{k[\gamma],a,b} = b]$ a *path* from a to b if

$Q(v_{i,a,b}, v_{i+1,a,b}) > 0$ for all $0 \le i < k[\gamma]$. Then let $\Gamma_{a,b}$ be the collection of all paths from a to b. Call a function F from paths to $[0,1]$ a *flow* if $\sum_{\gamma\in\Gamma_{a,b}} F[\gamma] = 1$ for all a, b. For a path $\gamma \in \Gamma_{a,b}$, we will label its initial and final vertices by $i(\gamma) = a$, $o(\gamma) = b$.

The purpose of these definitions is to provide a way to compare the functionals described in formula (5.2). If there exists a family of measures $\{\mathbb{P}_x\}_{x\in\widehat{\Theta}}$ so that the associated linear extensions given by formula (5.4) satisfy

$$L_\nu(f) \le C_L L_\mu(\widehat{f}),$$
$$\mathcal{E}_K(\widehat{f}, \widehat{f}) \le C_{\mathcal{E}} \mathcal{E}_Q(f, f),$$

then the variational characterization of α given in formula (5.3) implies

$$\alpha(Q) \ge \frac{1}{C_L C_{\mathcal{E}}} \alpha(K). \tag{5.5}$$

This is the motivation for Theorem 4 and Lemma 2 of [30]. Theorem 4 of [30] may be restated as the following.

THEOREM 4 (Comparison of Dirichlet forms for general chains). *Let K, Q be the kernels of two reversible Markov chains. Assume that K has stationary measure μ on state space $\widehat{\Theta}$ while Q has stationary measure ν on state space $\Theta \subset \widehat{\Theta}$. Fix flow F, distributions $\mathbb{P}_x$ and couplings $\mathbb{P}_{x,y}$ as in the notation in Definition 5.5 above. Then for any function f on Θ and the linear extension $\hat{f}$ of f on $\widehat{\Theta}$ given by formula (5.4),*

$$\mathcal{E}_K(\widehat{f}, \widehat{f}) \le \mathcal{A}\mathcal{E}_Q(f, f),$$

where

$$\begin{aligned}\mathcal{A} = \sup_{Q(q,r)>0} \frac{1}{Q(q,r)\nu(q)} \Bigg(& \sum_{\gamma \ni (q,r)} F[\gamma] k[\gamma] K\big(i(\gamma), o(\gamma)\big) \mu\big(i(\gamma)\big) \\ & + 2 \sum_{\gamma \ni (q,r)} k[\gamma] F[\gamma] \sum_{y \in \widehat{\Theta}\setminus\Theta} \mathbb{P}_y\big[o(\gamma)\big] K\big(i(\gamma), y\big) \mu\big(i(\gamma)\big) \\ & + \sum_{\gamma \ni (q,r)} k[\gamma] F[\gamma] \sum_{x,y \in \widehat{\Theta}\setminus\Theta : K(x,y)>0} \mathbb{P}_{x,y}\big[i(\gamma), o(\gamma)\big] K(x,y) \mu(x) \Bigg).\end{aligned}$$

Lemma 2 of [30] may be restated as the following.

LEMMA 5.6 (Comparison of variance and log-Sobolev constants). *Let μ be a measure on $\widehat{\Theta}$ and ν be a measure on $\Theta \subset \widehat{\Theta}$. Let $\tilde{C} = \sup_{y\in\Omega} \frac{\nu(y)}{\mu(y)}$. Then for any function f on Θ and linear extension $\hat{f}$ of f on $\widehat{\Theta}$,*

$$V_\nu(f) \le \tilde{C} V_\mu(\widehat{f}),$$
$$L_\nu(f) \le \tilde{C} L_\mu(\widehat{f}).$$

5.2. *Bounds on KCIP.* Next, we prove our results. Fix $1 \le k \le \frac{n}{2}$ and denote by $\alpha_{n,k}$ the log-Sobolev constant of $Q_{n,k}$. Denote by $Q^{\mathrm{SE}}_{n,k}$ and $\alpha^{\mathrm{SE}}_{n,k}$ the kernel and log-Sobolev constant associated with the simple exclusion process with k particles on $\Lambda(L,d)$. We then have the following.

LEMMA 5.7 (Comparison of log-Sobolev constants). *There exist* $N = N(c,k,d) < \infty$ *and* $0 < C = C(c,k,d) < \infty$ *so that* $n > N$ *implies*

$$\alpha_{n,k} \ge C\frac{1}{n}\alpha^{\mathrm{SE}}_{n,k}. \tag{5.6}$$

PROOF. Our proof consists of comparing a sequence of very similar Markov chains, beginning with the trace of the KCIP and ending with the simple exclusion process. The bulk of our argument goes through repeated application of Theorem 4, Lemma 5.6 and inequality (5.5).

Recall that the state space of $Q_{n,k}$ is $\Omega_{n,k} = \Omega_k$ as defined in formula (2.1), while the state space of $Q^{\mathrm{SE}}_{n,k}$ is $\Omega^{\mathrm{SE}}_{n,k}$ as defined in formula (5.1). By the standard "birthday problem" computation,

$$\prod_{i=1}^{k}\left(1-(i-1)\frac{(2d+1)}{n}\right) \le \frac{|\Omega_{n,k}|}{|\Omega^{\mathrm{SE}}_{n,k}|} \le 1.$$

Thus, for any fixed $k \in \mathbb{N}$,

$$1 - o(1) \le \frac{|\Omega_{n,k}|}{|\Omega^{\mathrm{SE}}_{n,k}|} \le 1 \tag{5.7}$$

as n goes to infinity. Since the stationary distributions $\pi_{n,k}$ and $\pi^{\mathrm{SE}}_{n,k}$ of $Q_{n,k}$ and $Q^{\mathrm{SE}}_{n,k}$, respectively, are uniform on $\Omega_{n,k}$ and $\Omega^{\mathrm{SE}}_{n,k}$, respectively, inequality (5.7) implies that

$$1 \le \frac{\pi_{n,k}(x)}{\pi^{\mathrm{SE}}_{n,k}(x)} \le 1 + o(1) \tag{5.8}$$

uniformly in $x \in \Omega_{n,k} \subset \Omega^{\mathrm{SE}}_{n,k}$.

Next, we define a less-lazy version of $Q_{n,k}$. Note that

$$Q_{n,k}(x,x) \ge 1 - \frac{2cdk}{n^2},$$

and so for n sufficiently large we can define a less-lazy version $Q^{\mathrm{NL}}_{n,k}$ of $Q_{n,k}$ by

$$Q_{n,k} = \left(1-\frac{c}{n}\right)\mathrm{Id} + \frac{c}{n}Q^{\mathrm{NL}}_{n,k},$$

where Id is the identity kernel and $Q^{\mathrm{NL}}_{n,k}$ is at least $\frac{1}{2}$-lazy itself. Since $Q^{\mathrm{NL}}_{n,k}$ is simply a less lazy version of $Q_{n,k}$, it is immediate that the associated Dirichlet forms $\mathcal{E}_{Q^{\mathrm{NL}}_{n,k}}$ and $\mathcal{E}_{Q_{n,k}}$ satisfy

$$\mathcal{E}_{Q_{n,k}}(f,f) \geq \frac{c}{n}\mathcal{E}_{Q^{\mathrm{NL}}_{n,k}}(f,f)$$

for all $f : \Omega_{n,k} \to \mathbb{R}$. Applying Lemma 5.6 to inequality (5.5), this bound implies the following inequality for the log-Sobolev constant $\alpha^{\mathrm{NL}}_{n,k}$ of $Q^{\mathrm{NL}}_{n,k}$:

$$\alpha_{n,k} \geq \frac{c}{n}\alpha^{\mathrm{NL}}_{n,k}. \tag{5.9}$$

Thus, to prove inequality (5.6) it is sufficient to relate $Q^{\mathrm{NL}}_{n,k}$ and $Q^{\mathrm{SE}}_{n,k}$.

Next, we define $Q^{\mathrm{MH-SE}}_{n,k}$ to be the Metropolis–Hasting chain with proposal chain equal to $Q^{\mathrm{SE}}_{n,k}$ and target measure $\pi_{n,k} = \mathrm{Unif}(\Omega_{n,k})$. If $x \neq y \in \Omega_{n,k}$ satisfy $Q^{\mathrm{MH-SE}}_{n,k}(x,y) > 0$, then x and y differ at exactly two vertices u, v, with $x[v] = y[u] = 0$, $x[u] = y[v] = 1$. Let $\phi_0 = 0$ and inductively let $\phi_{i+1} = \inf\{t > \phi_i : X_t \neq X_{\phi_i}\}$ be the successive times at which the KCIP changes; we calculate

$$\begin{aligned}
Q^{\mathrm{NL}}_{n,k}(x,y) &\geq \frac{n}{c}Q_{n,k}(x,y)\\
&\geq \frac{n}{c}\mathbb{P}\big[X_{\phi_1}[v] = 1 | X_0 = x\big]\mathbb{P}\big[X_{\phi_2}[u] = 0 | X_0 = x, X_{\phi_1}[v] = 1\big]\\
&\geq \frac{n}{c}\frac{c}{n^2}\frac{\frac{1}{n}(1-\frac{c}{n})}{\frac{1}{n}(1-\frac{c}{n}) + \frac{2cd(k+1)}{n^2}} \qquad (5.10)\\
&= \frac{1}{n}\left(\frac{n-c}{n-c+2cd(k+1)}\right)\\
&\geq \frac{1}{2}\left(\frac{n-c}{n-c+2cd(k+1)}\right)Q^{\mathrm{MH-SE}}_{n,k}(x,y).
\end{aligned}$$

To see the second and third lines, consider starting at KCIP at X_0 and calculating the probability that X_{ϕ_1} is obtained from X_0 by changing the label of v from 0 to 1, and that X_{ϕ_2} is obtained from X_{ϕ_1} by changing the label of u from 1 to 0; these probabilities are at least $\frac{c}{n^2}$ and $\frac{\frac{1}{n}(1-\frac{c}{n})}{\frac{1}{n}(1-\frac{c}{n})+\frac{2cd(k+1)}{n^2}}$ respectively. By the same short argument immediately preceding inequality (5.9), inequality (5.10) implies that for all n sufficiently large the log-Sobolev constant $\alpha^{\mathrm{MH-SE}}_{n,k}$ of $Q^{\mathrm{MH-SE}}_{n,k}$ satisfies

$$\alpha^{\mathrm{NL}}_{n,k} \geq \frac{1}{4}\alpha^{\mathrm{MH-SE}}_{n,k}. \tag{5.11}$$

In light of this inequality and inequality (5.9), to prove inequality (5.6) it is enough to relate $Q^{\mathrm{MH-SE}}_{n,k}$ and $Q^{\mathrm{SE}}_{n,k}$.

We now give the main comparison argument, using Theorem 4. Using the notation of Theorem 4, we will compare kernels $Q = Q_{n,k}^{\mathrm{MH-SE}}$ and $K = Q_{n,k}^{\mathrm{SE}}$ on state spaces $\Theta = \Omega_{n,k}$ and $\widehat{\Theta} = \Omega_{n,k}^{\mathrm{SE}}$. Both of these kernels have stationary distributions that are uniform on their respective state spaces. To define the flows, distributions and couplings required by Theorem 4, we need slightly more notation. Define the graphs $G_{n,k}^{\mathrm{MH-SE}}$ and $G_{n,k}^{\mathrm{SE}}$ to have vertices

$$V\big(G_{n,k}^{\mathrm{MH-SE}}\big) = \Omega_{n,k},$$
$$V\big(G_{n,k}^{\mathrm{SE}}\big) = \Omega_{n,k}^{\mathrm{SE}}$$

and edges

$$E\big(G_{n,k}^{\mathrm{MH-SE}}\big) = \big\{(x,y) \in G_{n,k}^{\mathrm{MH-SE}} : Q_{n,k}^{\mathrm{MH-SE}}(x,y) > 0\big\},$$
$$E\big(G_{n,k}^{\mathrm{SE}}\big) = \big\{(x,y) \in G_{n,k}^{\mathrm{SE}} : Q_{n,k}^{\mathrm{SE}}(x,y) > 0\big\}.$$

We denote by $d^{\mathrm{MH-SE}}$ and d^{SE} the usual graph distances on $G_{n,k}^{\mathrm{MH-SE}}$ and $G_{n,k}^{\mathrm{SE}}$, respectively. Next, define the distributions $\{\mathbb{P}_x\}_{x \in \Omega_{n,k}^{\mathrm{SE}}}$ on $\Omega_{n,k}$ by $\mathbb{P}_x[\cdot] = \delta_x(\cdot)$ if $x \in \Omega_{n,k}$, and

$$\mathbb{P}_x[\cdot] = \mathrm{Unif}\Big(\Big\{y \in \Omega_{n,k} : d^{\mathrm{SE}}(x,y) = \min_{z \in \Omega_{n,k}} d^{\mathrm{SE}}(x,z)\Big\}\Big)$$

otherwise. We define the couplings $\{\mathbb{P}_{x,y}\}_{x,y \in \Omega_{n,k}^{\mathrm{SE}}, Q_{n,k}^{\mathrm{SE}}(x,y)>0}$ to be the independent couplings $\mathbb{P}_{x,y}[a,b] = \mathbb{P}_x[a]\mathbb{P}_y[b]$. Finally, for pairs (a,b) satisfying $\sum_{x,y \in \Omega_{n,k}^{\mathrm{SE}}, Q_{n,k}^{\mathrm{SE}}(x,y)>0} \mathbb{P}_{x,y}[a,b] > 0$, we define the flow F on $\Gamma_{a,b}$ to be uniform on all minimum-length paths in $G^{\mathrm{MH-SE}}$ from a to b.

We now show that the constant $\mathcal{A}$ in Theorem 4 that the above choices yield is uniformly bounded in n. Recall that $\frac{|\Theta|}{|\widehat{\Theta}|} = 1 + o(1)$ and $\frac{Q^{\mathrm{MH-SE}}(x_1,y_1)}{Q^{\mathrm{SE}}(x_2,y_2)}$ is either 0 or 1 when it is defined, and so the constant $\mathcal{A}$ can be bounded by

$$\mathcal{A} = \big(1 + o(1)\big) \sup_{Q(q,r)>0} \Bigg(\sum_{\gamma \ni (q,r)} F[\gamma]k[\gamma] + 2 \sum_{\gamma \ni (q,r)} k[\gamma]F[\gamma] \sum_{y \in \widehat{\Theta} \setminus \Theta} \mathbb{P}_y\big[o(\gamma)\big]$$
$$+ \sum_{\gamma \ni (q,r)} k[\gamma]F[\gamma] \sum_{x,y \in \widehat{\Theta} \setminus \Theta : K(x,y)>0} \mathbb{P}_{x,y}\big[i(\gamma), o(\gamma)\big] \Bigg).$$

To show that $\mathcal{A}$ is uniformly bounded in n, it is enough to check that all of the probabilities, couplings and paths that we have defined are *local* in the sense that any particular probability, coupling or path involve only points that are a bounded distance from each other, uniformly in n. Since the details of the bounds are not important to us, we give very loose bounds.

$\mathbb{P}_x$ is concentrated on points y with $d^{\mathrm{SE}}(x, y) \le 2k^2$. $\mathbb{P}_{x,y}$ is defined only on pairs (x, y) satisfying $d^{\mathrm{SE}}(x, y) = 1$, and thus is supported on pairs (a, b) satisfying

$$d^{\mathrm{SE}}(a, b) \le 4k^2 + 1. \tag{5.12}$$

For $x, y \in \Omega_{n,k}$, let

$$C_{x,y} = \max\{|u - v| : u, v \in \Lambda(L, d), x[u] + y[u] \ge 1, x[v] + y[v] \ge 1\}$$

be the maximum distance between any particles in x or y. For any $0 < C < \infty$, let

$$R_{C,n} = \max\{d^{\mathrm{MH-SE}}(x, y) : C_{x,y} \le C\}.$$

Since $G^{\mathrm{MH-SE}}$ is vertex-transitive and is connected for n sufficiently large, we have for all $n > N(C, k, d)$ sufficiently large that

$$R_{C,n} \le \mathcal{R}_C$$

for some constant $\mathcal{R}_C$ that does not depend on n. Thus, for all $n > N(C, k, d)$ sufficiently large and all x, y with $C_{x,y} \le C$,

$$d^{\mathrm{MH-SE}}(x, y) \le \mathcal{R}_C. \tag{5.13}$$

If $d^{\mathrm{SE}}(x, y) \le C$, then there exist k vertices $v_1, \dots, v_k \in \Lambda(L, d)$ that cover the particles of x and y in the following sense:

$$\{v \in \Lambda(L, d) : x[v] + y[v] \ge 1\} \subset \bigcup_{i=1}^{k} \mathcal{B}_{2C}(v_i).$$

By taking larger balls, this can be turned into a disjoint cover: there exist $1 \le m \le k$ vertices $u_1, \dots, u_m$ so that

$$\{v \in \Lambda(L, d) : x[v] + y[v] \ge 1\} \subset \bigcup_{i=1}^{m} \mathcal{B}_{3k(C+\mathcal{R}_C)}(u_i)$$

with $\mathcal{B}_{3k(C+\mathcal{R}_C)}(u_i) \cap \mathcal{B}_{3k(C+\mathcal{R}_C)}(u_j) = \varnothing$ for all $i \ne j$. By the definition of $\mathcal{R}_C$, no minimal-length path from x to y can have any particles outside of the cover $\bigcup_{i=1}^{m} \mathcal{B}_{3k(C+\mathcal{R}_C)}(u_i)$; thus, by inequality (5.13), for all $x, y \in \Omega_{n,k}$ with $d^{\mathrm{SE}}(x, y) \le C$ we have that

$$d^{\mathrm{MH-SE}}(x, y) \le R_{3k(C+\mathcal{R}_C)}.$$

Combining this bound with inequality (5.12), we conclude that all paths γ with $F(\gamma) > 0$ have at most $\mathcal{R}_{3k(4k^2+1+\mathcal{R}_{4k^2+1})}$ points. Since balls of radius ℓ in the graph $\Lambda(L, d)$ have at most $(2\ell+1)^d$ vertices in them, balls of radius ℓ in $G_{n,k}^{\mathrm{MH-SE}}$ have at most $(2\ell + k)^{kd}$ vertices in them. Thus, at most $(2\mathcal{R}_{3k(4k^2+1+\mathcal{R}_{4k^2+1})} + k)^{kd} + k$ paths in $G_{n,k}^{\mathrm{MH-SE}}$ with positive support can pass through any given edge.

Most importantly, all of these bounds depend only on k and d; they are uniform in n sufficiently large. Combining this observation with inequality (5.8) and applying Theorem 4, this implies

$$\mathcal{E}_{Q_{n,k}^{\mathrm{SE}}}(\widehat{f},\widehat{f}) \le \mathcal{A}_{k,d}\mathcal{E}_{Q_{n,k}^{\mathrm{MH-SE}}}(f,f) \tag{5.14}$$

for some constant $\mathcal{A}_{k,d}$ that depends only on k and d. Thus, applying inequality (5.7) and Lemma 5.6 to inequality (5.5), we conclude that there exists some constant $C'_{k,d}$ so that

$$\alpha_{n,k}^{\mathrm{MH-SE}} \ge C'_{k,d}\alpha_{n,k}^{\mathrm{SE}}.$$

Combining this bound with inequalities (5.9) and (5.11) completes the proof of inequality (5.6). □

Finally, we prove Lemma 5.1:

PROOF OF LEMMA 5.1. Translating the main result of [32] into our discrete-time setting, we have

$$\alpha_{n,k}^{\mathrm{SE}} \ge C_d n^{-1-\frac{2}{d}},$$

for some constant C_d that depends only on d. By Lemma 5.7, this implies

$$\alpha_{n,k} \ge C n^{-2-\frac{2}{d}}$$

for some constant $C = C(c,d,k)$ that does not depend on n. Applying inequality (3.3) of [11] yields the conclusion. □

6. Drift condition for V_t. Recall the process $V_t = \sum_{v\in\Lambda(L,d)} X_t[v]$ from formula (2.3). The graph G_t is the subgraph of $G = \Lambda(L,d)$ induced by the vertices $\{v \in \Lambda(L,d) : X_t[v] = 1\}$, with vertices $V(G_t)$ and edges $E(G_t)$. Let $\mathcal{F}_t$ denote the σ-algebra generated by the random variables $\{X_s\}_{s\le t}$. The key result in this section is the following drift condition on $\{V_s\}_{s\ge t}$.

THEOREM 6.1. *There exists some constant $0 < \varepsilon_0 = \varepsilon_0(c,d)$ independent of n so that for all $0 < \varepsilon < \varepsilon_0$, there exist constants $C_G = C_G(\varepsilon,c,d) < \infty$, $\alpha = \alpha(\varepsilon,c,d)$ and $N = N(\varepsilon,c,d)$ so that, for all $k \in \mathbb{N}$ and all $n > N$,*

$$\mathbb{E}[V_{t+k\varepsilon n^3}|\mathcal{F}_t] \le (1-\alpha)^k V_t + C_G. \tag{6.1}$$

Besides a small number of definitions that are explicitly recalled, Theorem 6.1 is the only part of Section 6 used in the remaining sections. As the proof of Theorem 6.1 is somewhat long, we give an outline:

1. We show that the number of particles V_t is generally close to the number of connected components in G_t, and so it is enough to bound the latter quantity (see Lemma 6.2).

2. By embedding a coalescent process into the KCIP, we show that the number of "collisions" between components of the KCIP over the time interval $\{t, t+1, \ldots, t+\varepsilon n^3\}$ is on the order of the number of connected components in G_t (see Lemma 6.15).
3. By direct computation, we show that collisions involving components of size 1 will decrease a certain biased count of the number of connected components G_t (see Definition 6.3 and Lemma 6.6) while other collisions will not increase this observable by much (see Lemma 6.8).
4. By an argument based on bounding the influence of faraway points, we show that a substantial fraction of collisions occur between components of size 1 (see Lemma 6.9).
5. Steps (2)–(4) above will yield that a positive fraction of connected components of G_t will be involved in a "good" collision over a reasonable time scale, and this leads to a contraction estimate on a biased count of the number of components (see Lemma 6.11). By the observation made in (1) above, this leads to our final contraction estimate on V_t.

Recall from Section 4 that for a vertex u, $\mathrm{Comp}_t(u)$ is the collection of vertices contained in the same connected component as u in G_t. Recall also that Y_t denotes the number of connected components in G_t. Define the *number of excess particles*

$$\delta_s = V_s - Y_s. \tag{6.2}$$

The next lemma compares the number of particles to the number of components.

LEMMA 6.2. *For $n > 2c$ and for all $s, t \geq 0$*

$$\mathbb{E}[\delta_{t+s}|\mathcal{F}_t] \leq \left(1 - \frac{1}{2n}\right)^s \delta_t + 4cd. \tag{6.3}$$

PROOF. We first show that, for $k \in \mathbb{N}$,

$$\mathbb{E}[\delta_{t+s}\mathbf{1}_{\sup_{t\leq i\leq t+s} V_i \leq k}|\mathcal{F}_t] \leq \left(1 - \frac{1}{n}\left(1 - \frac{c}{n}\right)\right)^s \delta_t + \frac{2cdk}{n}\frac{1}{1-\frac{c}{n}}. \tag{6.4}$$

Indeed, for $v \in \Lambda(L,d)$, let $N_s^{\mathrm{adj}}(v)$ be the number of components of G_s that are adjacent to v. Define

$$\begin{aligned} A_s &= \{v \in \Lambda(L,d) : X_s[v] = 1, |\mathrm{Comp}_s(v)| > 1\}, \\ B_s &= \{v \in \Lambda(L,d) : X_s[v] = 0, N_s^{\mathrm{adj}}(v) > 0\}, \\ D_s &= \Lambda(L,d) \setminus (A_s \cup B_s). \end{aligned}$$

Then we have

$$\begin{aligned} \mathbb{E}[\delta_{s+1}|\mathcal{F}_s] = {} & \mathbb{E}[\delta_{s+1}|\mathcal{F}_s, v_s \in A_s]\mathbb{P}[v_s \in A_s|\mathcal{F}_s] \\ & + \mathbb{E}[\delta_{s+1}|\mathcal{F}_s, v_s \in B_s]\mathbb{P}[v_s \in B_s|\mathcal{F}_s] \end{aligned}$$

$$
\begin{aligned}
&\quad + \mathbb{E}[\delta_{s+1}|\mathcal{F}_s, v_s \in D_s]\mathbb{P}[v_s \in D_s|\mathcal{F}_s] \\
&\le \left(\delta_s - \left(1 - \frac{c}{n}\right)\right)\mathbb{P}[v_s \in A_s|\mathcal{F}_s] \\
&\quad + \left(\delta_s + \frac{c}{n}\right)\mathbb{P}[v_s \in B_s|\mathcal{F}_s] + \delta_s\mathbb{P}[v_s \in D_s|\mathcal{F}_s] \\
&= \delta_s\mathbb{P}[v_s \in A_s \cup B_s \cup D_s|\mathcal{F}_s] \\
&\quad - \left(1 - \frac{c}{n}\right)\mathbb{P}[v_s \in A_s|\mathcal{F}_s] + \frac{c}{n}\mathbb{P}[v_s \in B_s|\mathcal{F}_s] \\
&\le \delta_s - \frac{\delta_s}{n}\left(1 - \frac{c}{n}\right) + \frac{2cdV_s}{n^2}.
\end{aligned} \tag{6.5}
$$

This can be iterated to give:

$$
\begin{aligned}
&\mathbb{E}\big[(\delta_{t+s})\mathbf{1}_{\sup_{t\le i\le t+s} V_i \le k}|\mathcal{F}_t\big] \\
&\quad = \mathbb{E}\big[\mathbb{E}\big[(\delta_{t+s})\mathbf{1}_{\sup_{t\le i\le t+s} V_i \le k}|\mathcal{F}_{t+s-1}\big]|\mathcal{F}_t\big] \\
&\quad \le \mathbb{E}\left[\left(\left(1 - \frac{1}{n}\left(1 - \frac{c}{n}\right)\right)\delta_{t+s-1} + \frac{2cdV_{t+s-1}}{n^2}\right)\mathbf{1}_{\sup_{t\le i\le t+s} V_i \le k}\middle|\mathcal{F}_t\right] \\
&\quad \le \mathbb{E}\left[\left(\left(1 - \frac{1}{n}\left(1 - \frac{c}{n}\right)\right)\delta_{t+s-1} + \frac{2cdk}{n^2}\right)\mathbf{1}_{\sup_{t\le i\le t+s} V_i \le k}\middle|\mathcal{F}_t\right] \\
&\quad \le \cdots \\
&\quad \le \left(1 - \frac{1}{n}\left(1 - \frac{c}{n}\right)\right)^s \delta_t + \frac{2cdk}{n}\frac{1}{1 - \frac{c}{n}},
\end{aligned}
$$

which is inequality (6.4). Since $V_s \le n$, this implies inequality (6.3) for $n > 2c$ and the proof is completed. □

6.1. *Corrected number of components.* Our next goal is to obtain a bound on $\mathbb{E}[Y_{t+s}|\mathcal{F}_t]$ in terms of Y_t over certain time intervals. To this end, we digress briefly and introduce a new object called the "corrected number of components" of a graph.

DEFINITION 6.3. For a graph H, we define a Markov chain $\{H_i\}_{i\ge 0}$ with absorbing states as follows. Set $H_0 = H$. For $i \ge 0$, if all components of H_i are size 1, set $H_{i+1} = H_i$. Otherwise, select uniformly at random a vertex v_i in H_i that also has at least one neighbour in H_i and set $H_{i+1} = H_i \setminus \{v_i\}$. We then define $\mathcal{N}_H = \mathbb{E}[\lim_{i\to\infty}|H_i|]$. Since $H_{i+1} = H_i$ for all $i > |H|$, this (random) limit always exists. See Figure 3 for an illustration of the evolution of H_t for an initial graph H_0 with five vertices.

DEFINITION 6.4. For $k \in \mathbb{N}$, set

$$\mathcal{N}_k = \sup\{\mathcal{N}_H : H \subset \Lambda(L, d), H \text{ connected}, |H| \le k\}.$$

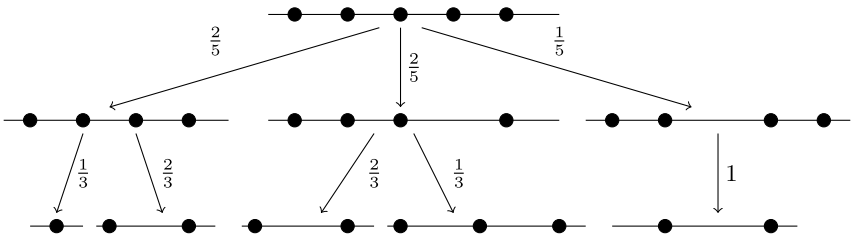

FIG. 3. *Evolution of the component sizes of H_t: These are the transition probabilities associated with the $\{H_t\}_{t\in\mathbb{N}}$ process, with H_0 a 5-vertex graph.*

We make some initial observations. By simple case checking, we can show that $\mathcal{N}_1 = \mathcal{N}_2 = 1, \mathcal{N}_3 = \frac{4}{3}, \mathcal{N}_4 = \frac{7}{4}$. We clearly also have

$$\mathcal{N}_k \le k - 1. \tag{6.6}$$

Next, we show that $\mathcal{N}_H$ can never be too small.

LEMMA 6.5. *For any subgraph $H \subset \Lambda(L, d)$, we have*

$$\mathcal{N}_H \ge \frac{|H|}{2d+1}.$$

PROOF. Let $\{H_i\}_{i\ge 0}$ be the Markov process given in Definition 6.3, and let H_∞ be its limit. We have

$$\mathcal{N}_H = \mathbb{E}\Big[\sum_{v\in H} \mathbf{1}_{v\in H_\infty}\Big] = \sum_{v\in H} \mathbb{P}[v \in H_\infty]. \tag{6.7}$$

However, for any particular vertex $v \in H$, $v \in H_\infty$ as long as it is the last vertex among its neighbours to be selected as an update vertex v_i. Since v has at most $2d$ neighbours in H,

$$\mathbb{P}[v \in H_\infty] \ge \frac{1}{2d+1}.$$

Combining this with formula (6.7),

$$\mathcal{N}_H \ge \frac{|H|}{2d+1},$$

completing the proof. □

For $1 \le k \le 2d$, let $H(k)$ be the *star graph* with k leaves (see Figure 4 for a star graph with $d = 2$, $k = 4$).

FIG. 4. *Star graph.*

For the purposes of the following lemma, we denote the vertex set of this graph by $V(H(k)) = \{1, 2, \dots, k+1\}$ and the edge set by $E(V(k)) = \{(i, k+1) : 1 \le i \le k\}$. We give basic bounds on how $\mathcal{N}_{H(k)}$ depends on small changes to the subgraph H.

LEMMA 6.6 (Shrinking star graphs). *For $\mathcal{N}_H$ as in Definition 6.3,*

$$\mathcal{N}_{H(k)} = \frac{k}{2} + \frac{1}{k+1}. \tag{6.8}$$

PROOF. It is easy to check that $\mathcal{N}_{H(1)} = 1$ and $\mathcal{N}_{H(2)} = \mathcal{N}_3 = \frac{4}{3}$. Let $\{H_i\}_{i\in\mathbb{N}}$ be the Markov chain described in Definition (6.3), and let $\{v_i\}_{i\in\mathbb{N}}$ be the associated sequence of vertices. The quantity $\mathcal{N}_{H(k)}$ satisfies the recurrence:

$$\begin{aligned}\mathcal{N}_{H(k)} &= k\mathbb{P}[v_0 = k+1] + \mathbb{P}[v_0 \neq k+1]\mathcal{N}_{H(k-1)} \\ &= \frac{k}{k+1} + \frac{k}{k+1}\mathcal{N}_{H(k-1)}.\end{aligned} \tag{6.9}$$

Iterating, for all $0 \le q \le k-3$ we have

$$\mathcal{N}_{H(k)} = \frac{1}{k+1}\big(k + (k-1) + \cdots + (k-q)\big) + \frac{k-q}{k+1}\mathcal{N}_{H(k-q-1)}.$$

Setting $q = k-3$ gives formula (6.8). □

More generally, we have the following.

LEMMA 6.7. *For any graph G and all subgraphs $H \subset G$ and all vertices $v \in G$,*

$$\mathcal{N}_{H\cup\{v\}} \le \mathcal{N}_H + 1.$$

This bound holds regardless of whether or not H is connected, and regardless of whether or not v is adjacent to any vertex in H.

PROOF. We prove this by induction on $m = |H|$. For $|H| \in \{1, 2\}$ this is clear by direct computation. Assume that $\mathcal{N}_{H\cup\{v\}} \le \mathcal{N}_H + 1$ holds for all $|H| \le m$. Fix H with $|H| = m+1$, define $N(H)$ to be the vertices of H with at least one neighbour in H and, by the same argument as in recurrence (6.9),

$$\begin{aligned}\mathcal{N}_{H\cup\{v\}} &= \frac{1}{|N(H\cup\{v\})|}\sum_{u\in N(H\cup\{v\})}\mathcal{N}_{H\cup\{v\}\setminus\{u\}} \\ &\le \frac{1}{|N(H\cup\{v\})|}\mathcal{N}_H\mathbf{1}_{v\in N(H\cup\{v\})} \\ &\quad + \frac{1}{|N(H\cup\{v\})|}\sum_{u\in N(H\cup\{v\})\setminus\{v\}}(\mathcal{N}_{H\setminus\{u\}} + 1) \\ &\le \mathcal{N}_H + 1,\end{aligned}$$

where the induction hypothesis is used in the second line. □

6.2. *A corrected version of of* Y_t. In order to bound $Y_s - Y_t$, we introduce the process

$$\tilde{Y}_s = \sum_{H \in \mathrm{Comp}_s} \mathcal{N}_H. \tag{6.10}$$

We think of $\{\tilde{Y}_s\}_{s \geq t}$ as a "corrected" version of $\{Y_s\}_{s \geq t}$. In general, $\{\tilde{Y}_s\}_{s \geq t}$ is easier to work with than $\{Y_s\}_{s \geq t}$. One reason is the martingale-like property

$$\mathbb{E}\big[(\tilde{Y}_{t+1} - \tilde{Y}_t)\mathbf{1}_{p_t > \frac{c}{n}} | \mathcal{F}_t\big] = 0,$$

which does not hold for $\{Y_s\}_{s \geq t}$ [here p_t is as defined in (1.1)]. In addition, $\{\tilde{Y}_s\}_{s \geq t}$ is much better-behaved than $\{Y_s\}_{s \geq t}$ over short time intervals, especially when the KCIP is far from equilibrium. Using (6.6), it can be verified that

$$Y_s \leq \tilde{Y}_s \leq Y_s + \delta_s. \tag{6.11}$$

Since δ_s is often small (see Lemma 6.2), $\tilde{Y}_s$ is a good proxy for the quantity Y_s.

Next, we show that the corrected number of components $\{\tilde{Y}_s\}_{s \geq t}$ does not grow too quickly over a moderate time period.

LEMMA 6.8. *Fix* $0 < \varepsilon \leq \frac{1}{96c^2(d+1)^3}$. *Then for all* $0 \leq s \leq \varepsilon n^3$ *and all* $t \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}[\tilde{Y}_{t+s} | \mathcal{F}_t] \leq \tilde{Y}_t \left(1 + \frac{96c^2(d+1)^3}{n^3}\left(8n\left(1 - \frac{1}{8n}\right)^s + s\right)\right) + \frac{24cd}{n}\delta_t \\ + \left(1 - \frac{1}{8n}\right)^s \delta_t.\end{aligned} \tag{6.12}$$

PROOF. Assume without loss of generality that $t = 0$ and define

$$f(x) = \frac{3cd}{n^2} \sum_{u=0}^{x} \mathbb{E}[\delta_u | \mathcal{F}_0].$$

By Lemma 6.7, $\mathbb{E}[\tilde{Y}_{s+1} | \mathcal{F}_s] \leq \tilde{Y}_s + 1$ for any time s. Inequality (6.6) provides further necessary (but not sufficient) conditions for $\tilde{Y}_{s+1} > \tilde{Y}_s$ to hold when the update variable p_s in representation (1.1) satisfies $p_s < \frac{c}{n}$: the vertex v_s in representation (1.1) must be adjacent to a component of size at least two in G_s, and furthermore we must have $X_s[v_s] = 0$. At time s, there are at most $3d\delta_s$ vertices satisfying this necessary condition, and so for $0 \leq s \leq \varepsilon n^3$,

$$\begin{aligned}\mathbb{E}[\tilde{Y}_s | \mathcal{F}_0] &= \mathbb{E}\big[\mathbb{E}[\tilde{Y}_s | \mathcal{F}_{s-1}] | \mathcal{F}_0\big] \\ &= \mathbb{E}\left[\mathbb{E}\left[\tilde{Y}_s | \mathcal{F}_{s-1}, \left\{p_{s-1} \leq \frac{c}{n}\right\}\right]\mathbb{P}\left[p_{s-1} \leq \frac{c}{n} \middle| \mathcal{F}_{s-1}\right]\right.\end{aligned}$$

$$
\begin{aligned}
&\quad + \mathbb{E}\Big[\tilde{Y}_s \big| \mathcal{F}_{s-1}, \Big\{p_{s-1} > \frac{c}{n}\Big\}\Big]\mathbb{P}\Big[p_{s-1} > \frac{c}{n}\big|\mathcal{F}_{s-1}\Big]\Big|\mathcal{F}_0\Big] \\
&\le \mathbb{E}\Big[\Big(\tilde{Y}_{s-1} + \frac{3d\delta_{s-1}}{n}\Big)\mathbb{P}\Big[p_{s-1} \le \frac{c}{n}\Big|\mathcal{F}_{s-1}\Big] \\
&\quad + \tilde{Y}_{s-1}\mathbb{P}\Big[p_{s-1} > \frac{c}{n}\big|\mathcal{F}_{s-1}\Big]\Big|\mathcal{F}_0\Big] \qquad (6.13)\\
&\le \mathbb{E}[\tilde{Y}_{s-1}|\mathcal{F}_0] + \frac{3cd}{n^2}\mathbb{E}[\delta_{s-1}|\mathcal{F}_0] \\
&\le \cdots \\
&\le \tilde{Y}_0 + \frac{3cd}{n^2}\sum_{u=0}^{s-1}\mathbb{E}[\delta_u|\mathcal{F}_0].
\end{aligned}
$$

Lemma 6.5 implies

$$
\tilde{Y}_s \ge \frac{V_s}{2d+1}. \qquad (6.14)
$$

Using this fact, and then inequality (6.5) followed by inequality (6.13), we find that for all n sufficiently large,

$$
\begin{aligned}
f(s) &\equiv \frac{3cd}{n^2}\sum_{u=0}^{s}\mathbb{E}[\delta_u|\mathcal{F}_0] \\
&= \frac{3cd}{n^2}\Big(\delta_0 + \sum_{u=1}^{s}\mathbb{E}\big[\mathbb{E}[\delta_u|\mathcal{F}_{u-1}]|\mathcal{F}_0\big]\Big) \\
&\le \frac{3cd}{n^2}\Big(\delta_0 + \sum_{u=1}^{s}\mathbb{E}\Big[\Big(1-\frac{1}{4n}\Big)\delta_{u-1} + \frac{2cdV_{u-1}}{n^2}\Big|\mathcal{F}_0\Big]\Big) \\
&\le \frac{3cd}{n^2}\Big(\delta_0 + \Big(1-\frac{1}{4n}\Big)\sum_{u=0}^{s-1}\mathbb{E}[\delta_u|\mathcal{F}_0] + \frac{4c(d+1)^2}{n^2}\sum_{u=0}^{s-1}\mathbb{E}[\tilde{Y}_u|\mathcal{F}_0]\Big) \qquad (6.15)\\
&\le \frac{3cd}{n^2}\Big(\delta_0 + \Big(1-\frac{1}{4n}\Big)\sum_{u=0}^{s-1}\mathbb{E}[\delta_u|\mathcal{F}_0] + \frac{4c(d+1)^2}{n^2}(s-1)\tilde{Y}_0 \\
&\quad + \frac{12c^2(d+1)^3}{n^4}\sum_{u=0}^{s-2}u\mathbb{E}[\delta_u|\mathcal{F}_0]\Big) \\
&\le \frac{3cd}{n^2}\delta_0 + \frac{12c^2(d+1)^3}{n^4}(s-1)\tilde{Y}_0 + \frac{3cd}{n^2}\Big(1-\frac{1}{8n}\Big)\sum_{u=0}^{s-1}\mathbb{E}[\delta_u|\mathcal{F}_0] \\
&= \frac{3cd}{n^2}\delta_0 + \frac{12c^2(d+1)^3}{n^4}(s-1)\tilde{Y}_0 + \Big(1-\frac{1}{8n}\Big)f(s-1).
\end{aligned}
$$

Inequality (6.13) is used in the fifth step of this bound and we use the hypothesis that $\varepsilon \le \frac{1}{96c^2(d+1)^3}$ in the sixth step. Using this calculation with the last line of inequality (6.13), we have

$$
\begin{aligned}
\mathbb{E}[\tilde{Y}_s|\mathcal{F}_0] &\le \tilde{Y}_0 + f(s-1) \\
&\le \tilde{Y}_0\biggl(1 + \frac{12c^2(d+1)^3}{n^4}(s-1)\biggr) + \frac{3cd}{n^2}\delta_0 + \biggl(1 - \frac{1}{8n}\biggr) f(s-2) \\
&\le \cdots \\
&\le \tilde{Y}_0\biggl(1 + \sum_{x=0}^{s-q} \frac{12c^2(d+1)^3}{n^4} x\biggl(1 - \frac{1}{8n}\biggr)^{s-x}\biggr) \\
&\quad + \frac{3cd}{n^2}\delta_0 \sum_{x=0}^{s-q}\biggl(1 - \frac{1}{8n}\biggr)^{x} + \biggl(1 - \frac{1}{8n}\biggr)^{s-q} f(q) \\
&\le \cdots \\
&\le \tilde{Y}_0\biggl(1 + \sum_{x=0}^{s} \frac{12c^2(d+1)^3}{n^4} x\biggl(1 - \frac{1}{8n}\biggr)^{s-x}\biggr) \\
&\quad + \frac{3cd}{n^2}\delta_0 \sum_{x=0}^{s}\biggl(1 - \frac{1}{8n}\biggr)^{x} + \biggl(1 - \frac{1}{8n}\biggr)^{s} f(0) \\
&\le \tilde{Y}_0\biggl(1 + \frac{96c^2(d+1)^3}{n^3}\biggl(8n\biggl(1 - \frac{1}{8n}\biggr)^{s} + s\biggr)\biggr) \\
&\quad + \frac{24cd}{n}\delta_0 + \biggl(1 - \frac{1}{8n}\biggr)^{s}\delta_0.
\end{aligned}
\tag{6.16}
$$

This completes the proof of inequality (6.12), and thus the lemma. □

6.3. *Typical component size involved in collisions.* In this subsection, we study the behaviour of a "typical" collision in the KCIP by looking at the behaviour of the KCIP only for vertices and times close to the collision. This approach of "zooming in" on a neighbourhood of a collision is illustrated in Figure 5.

Fix $t > 0$. Define the *number of collisions* of the KCIP to be

$$
\mathcal{C}_s = \bigl|\{t \le u < s : Y_{u+1} < Y_u\}\bigr|, \tag{6.17}
$$

and define the *set of collision times*

$$
\mathcal{T}_{\mathrm{col}}(s) = \{t \le u \le t + s : Y_{u+1} < Y_u\}.
$$

Recall the definition of the update variable v_s used in formula (1.1). For $u \in \mathcal{T}_{\mathrm{col}}(s)$, we define

$$
\mathcal{M}_{\mathrm{col}}(u) = \sup\bigl\{\bigl|\mathrm{Comp}_u(w)\bigr| : (v_u, w) \in E\bigr\}
$$

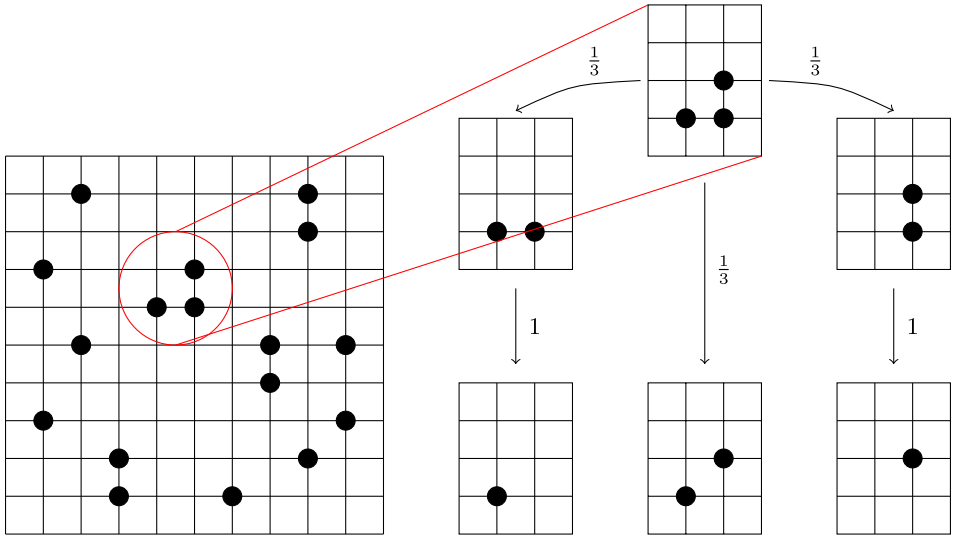

FIG. 5. *We study the KCIP dynamics in a small neighbourhood of a collision. The behaviour of the KCIP process outside of this neighbourhood has a bounded influence on the behaviour of the collision, effectively allowing us to ignore complicated configurations outside of the neighbourhood.*

to be the largest component involved in a collision; for $u \notin \mathcal{T}_{\mathrm{col}}(s)$ we set $\mathcal{M}_{\mathrm{col}}(u) = 0$. We first show that, given $u \in \mathcal{T}_{\mathrm{col}}(s)$, $\mathcal{M}_{\mathrm{col}}(u)$ is often equal to 1.

LEMMA 6.9 (Typical component size). *There exists $\delta = \delta(c,d) > 0$ independent of n so that for any $\varepsilon > 0$*

$$\begin{aligned}&\mathbb{E}\big[\big|\{t + n\log(n)^4 \le u \le t + \varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) = 1\}\big|\big|\mathcal{F}_t\big] \\ &\quad \ge (\delta - o(n^{-1}))\mathbb{E}\big[\big|\{t + n\log(n)^4 \le u \le t + \varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) \ne 0\}\big|\big|\mathcal{F}_t\big]\end{aligned} \tag{6.18}$$

uniformly in the initial configuration X_t.

REMARK 6.10. We briefly discuss why the conclusion obtained in Lemma 6.9 is plausible. Fix a time t and consider any vertex v and any neighbourhood $\mathcal{B}(v)$ whose size does not grow with n. It is easy to check that over any time interval of length T, with $n \ll T \ll n^2$, it is very likely that all vertices in $\mathcal{B}(v)$ have been updated and also that no particles have been added to $\mathcal{B}(v)$. On that very likely event, all components of $G_{t+T} \cap \mathcal{B}(v)$ are of size 1. In other words, after a burn-in time of length $O(n)$, most vertices are not close to any components of size greater than 1. The proof largely consists of checking that conditioning on a collision occurring at vertex v at times in the interval $(t + T, t + \varepsilon n^3)$ does not affect this conclusion too much.

PROOF OF LEMMA 6.9. Set $T = n\log(n)^4$. Let $\{p_s\}_{s\in\mathbb{N}}, \{v_s\}_{s\in\mathbb{N}}$ be the update variables used to define the dynamics of the KCIP as given in formula (1.1). For fixed $v \in \Lambda(L,d)$ and $t+T \le r \le t+\varepsilon n^3$, define the event

$$\mathcal{A}_{v,r}^{(\mathrm{coll})} = \{r \in \mathcal{T}_{\mathrm{col}}(\varepsilon n^3), v_r = v\}.$$

Thus, $\mathcal{A}_{v,r}^{(\mathrm{coll})}$ denotes the event that a collision occurs at time r and vertex v. We will show that, for any $v \in \Lambda(L,d)$, uniformly in $t+T \le r \le t+\varepsilon n^3$,

$$\mathbb{P}[\mathcal{M}_{\mathrm{col}}(r) = 1 | \mathcal{F}_t, \mathcal{A}_{v,r}^{(\mathrm{coll})}] \ge (\delta - o(n^{-1})) \tag{6.19}$$

for some $\delta > 0$. Inequality (6.19) will immediately yield (6.18). To see this, (6.19) implies that

$$\begin{aligned}
&\mathbb{E}[|\{t+T \le u \le t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) = 1\}||\mathcal{F}_t] \\
&\quad = \sum_{v\in\Lambda(L,d)} \sum_{r=t+T}^{t+\varepsilon n^3} \mathbb{P}[\mathcal{M}_{\mathrm{col}}(r) = 1 | \mathcal{F}_t, \mathcal{A}_{v,r}^{(\mathrm{coll})}]\mathbb{P}[\mathcal{A}_{v,r}^{(\mathrm{coll})}|\mathcal{F}_t] \\
&\quad \ge (\delta - o(n^{-1})) \sum_{v\in\Lambda(L,d)} \sum_{r=t+T}^{t+\varepsilon n^3} \mathbb{P}[\mathcal{A}_{v,r}^{(\mathrm{coll})}|\mathcal{F}_t] \\
&\quad = (\delta - o(n^{-1}))\mathbb{E}[|\{t+T \le u \le t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) \ne 0\}||\mathcal{F}_t],
\end{aligned}$$

yielding inequality (6.18). Thus, it suffices to show (6.19).

We begin with a simple bound on the probability that the number of particles in a small region ever goes up by more than a small number over any small time interval. For $n \in \mathbb{N}$ and $0 \le q \le 1$, denote by Binomial(n, q) a binomial random variable with n trials and success probability q. For any fixed $t \le r \le t+\varepsilon n^3$, $v \in \Lambda(L,d)$ and $\ell, k \in \mathbb{N}$, we have for $n > \sqrt{c(2\ell+1)^d}$,

$$\begin{aligned}
\mathbb{P}\left[\sum_{s=r+1}^{r+T} \mathbf{1}_{v_s\in\mathcal{B}_\ell(v)}\mathbf{1}_{p_s<\frac{c}{n}} \ge k\right] &\le \mathbb{P}\left[\mathrm{Binomial}\left(T, \frac{c(2\ell+1)^d}{n^2}\right) \ge k\right] \\
&\le \mathbb{P}\left[\mathrm{Binomial}\left(T, \frac{c(2\ell+1)^d}{n^2}\right) \ge 1\right]^k \\
&\le \left(\frac{c(2\ell+1)^d \log(n)^4}{n}\right)^k,
\end{aligned} \tag{6.20}$$

where the first inequality uses the simple bound $|\mathcal{B}_\ell(v)| \le (2\ell+1)^d$.

For $v \in \Lambda(L,d)$, $\ell \in \mathbb{N}$, and $t+T \le r \le t+\varepsilon n^3$, define the event

$$\mathcal{A}_{v,r,\ell}^{(\mathrm{sparse})} = \left\{\sum_{s=r-T}^{r-1} \mathbf{1}_{v_s\in\mathcal{B}_\ell(v)}\mathbf{1}_{p_s<\frac{c}{n}} < 8\right\}$$

and let $\mathcal{A}^{(\text{sparse})} = \bigcap_{t+T\le r\le t+\varepsilon n^3, v\in\Lambda(L,d)} \mathcal{A}^{(\text{sparse})}_{v,r,100}$. By inequality (6.20) and a union bound,

$$\mathbb{P}\big[\mathcal{A}^{(\text{sparse})}|\mathcal{F}_t\big] \ge 1 - O\bigg(n^4\frac{\log(n)^{32}}{n^8}\bigg) = 1 - o\big(n^{-3}\big). \tag{6.21}$$

From (6.21), we thus have

$$\begin{aligned}
&\mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}\big] \\
&\quad\ge \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big]\mathbb{P}\big[\mathcal{A}^{(\text{sparse})}_{v,r,100}|\mathcal{F}_t\big] \\
&\quad\ge \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big]\big(1 - o\big(n^{-3}\big)\big).
\end{aligned} \tag{6.22}$$

In light of (6.22), to show (6.19), it is enough to show

$$\mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big] \ge \big(\delta - o\big(n^{-1}\big)\big). \tag{6.23}$$

To this end, we proceed by defining two events of interest. For $v \in \Lambda(L,d)$, $t + T \le r \le t + \varepsilon n^3$, and some constant B_1 to be determined later, define the event

$$\mathcal{A}^{(\text{gap})}_{v,r} = \bigg\{ \sum_{u=r-B_1 n}^{r} \mathbf{1}_{p_u < \frac{c}{n}} \mathbf{1}_{v_u \in \mathcal{B}_{100}(v)} = 0 \bigg\}$$

that no vertices are added near v in the (short) time interval of length $B_1 n$ immediately before the collision. Finally, for $k \in \mathbb{N}$ and $t + T \le r \le t + \varepsilon n^3$, define the event

$$\mathcal{A}^{(\text{regular})}_{v,r,k} = \bigg\{ \sum_{u=r-T}^{r} \mathbf{1}_{v_u \in \mathcal{B}_{100}(v)} \le k \bigg\}$$

that at most k vertices are updated in a ball around v in the interval of length T before the collision.

Fix the constant $C_1 = 25(201)^d$. For any $t + T \le r \le t + \varepsilon n^3$, we have

$$\begin{aligned}
&\mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big] \\
&\quad\ge \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}, \mathcal{A}^{(\text{gap})}_{v,r}\big] \\
&\qquad\times \mathbb{P}\big[\mathcal{A}^{(\text{gap})}_{v,r}|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big] \\
&\quad\ge \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}, \mathcal{A}^{(\text{gap})}_{v,r}\big] \\
&\qquad\times \mathbb{P}\big[\mathcal{A}^{(\text{gap})}_{v,r}|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}, \mathcal{A}^{(\text{regular})}_{v,r,C_1\log(n)^4}\big] \\
&\qquad\times \mathbb{P}\big[\mathcal{A}^{(\text{regular})}_{v,r,C_1\log(n)^4}|\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}\big] \\
&\quad\equiv \mathrm{T}_1 \times \mathrm{T}_2 \times \mathrm{T}_3.
\end{aligned} \tag{6.24}$$

The remainder of the proof consists of obtaining a lower bound for each of the three factors $\mathtt{T}_1, \mathtt{T}_2$ and $\mathtt{T}_3$ in inequality (6.24). We begin by bounding $\mathtt{T}_3$. For a subgraph $H \subset \Lambda(L,d)$ and a configuration $X \in \{0,1\}^{\Lambda(L,d)}$, denote by $X|_H$ the restriction of X to H; that is, $X|_H \in \{0,1\}^H$ and satisfies $X|_H[w] = X[w]$ for all $w \in H$. Let

$$\psi_H = \big\{\psi_H(0), \psi_H(1), \ldots\big\} \equiv \{s \geq t : X_s|_H \neq X_{s+1}|_H\}$$

be the ordered sequence of times at which the restriction of $\{X_s\}_{s\in\mathbb{N}}$ to H changes. Finally, let $\mathcal{G}_{v,r}$ be the σ-algebra generated by the random variables $\{X_s|_{\mathcal{B}_{100}(v)}\}_{s\in\psi_{\mathcal{B}_{100}(v)}\cap\{r-T,\ldots,r-1\}}$ and $\psi_{\mathcal{B}_{100}(v)}$. For notational convenience, let $\mathcal{S}_t$ be shorthand for $\mathcal{F}_t$, $\mathcal{A}^{(\mathrm{coll})}_{v,r}$, $\mathcal{A}^{(\mathrm{sparse})}_{v,r,100}$ and denote by $\psi^c_{\mathcal{B}_{100}(v)}$ the complement of the set $\psi_{\mathcal{B}_{100}(v)}$. We have

$$\begin{aligned}
1 - \mathtt{T}_3 &= 1 - \mathbb{E}\big[\mathbb{P}\big[\mathcal{A}^{(\mathrm{regular})}_{v,r,C_1\log(n)^4}|\mathcal{S}_t, \mathcal{G}_{v,r}\big]|\mathcal{S}_t\big] \qquad (6.25)\\
&\leq \frac{1}{C_1\log(n)^4}\mathbb{E}\bigg[\mathbb{E}\bigg[\sum_{s=r-T}^{r-1}\mathbf{1}_{v_s\in\mathcal{B}_{100}(v)}\Big|\mathcal{S}_t, \mathcal{G}_{v,r}\bigg]\Big|\mathcal{S}_t\bigg]\\
&\leq \frac{1}{C_1\log(n)^4}\mathbb{E}\bigg[(201)^d + 16 \qquad (6.26)\\
&\quad + \mathbb{E}\bigg[\sum_{s\in\{r-T\leq u\leq r-1 : v_u\in\mathcal{B}_{100}(v)\}\cap\psi^c_{\mathcal{B}_{100}(v)}}\mathbf{1}_{v_s\in\mathcal{B}_{100}(v)}\Big|\mathcal{S}_t, \mathcal{G}_{v,r}\bigg]\Big|\mathcal{S}_t\bigg].
\end{aligned}$$

Inequality (6.25) is simply an application of Markov's inequality. To obtain (6.26), we split $\{s : v_s \in \mathcal{B}_{100}(v)\} \cap \{r-T, r-T+1, \ldots, r-1\}$ into two sets: $\{r-T,\ldots,r-1\}\cap\psi_{\mathcal{B}_{100}(v)}$ and everything else. Inequality (6.26) then follows from noting that the number $|\{r-T,\ldots,r-1\}\cap\psi_{\mathcal{B}_{100}(v)}|$ of times that a particle is added to or removed from $\mathcal{B}_{100}(v)$ between times $r-T$ and $r-1$ is, at most, the number of particles in that set at time $r-T$ plus twice the number that have been added between times $r-T$ and $r-1$. Thus, conditional on $\mathcal{A}^{\mathrm{sparse}}$, we have $|\{r-T,\ldots,r-1\}\cap\psi_{\mathcal{B}_{100}(v)}| \leq (201)^d + 16$.

Fix $s \in \{r-T,\ldots,r-1\}\cap\psi^c_{\mathcal{B}_{100}(v)}$. We claim that if $w_1, w_2 \in \Lambda(L,d)\setminus\mathcal{B}_{108}(v)$, then

$$\frac{\mathbb{P}[v_s = w_1|\mathcal{S}_t, \mathcal{G}_{v,r}]}{\mathbb{P}[v_s = w_2|\mathcal{S}_t, \mathcal{G}_{v,r}]} = 1, \qquad (6.27)$$

since updates to these two vertices cannot influence anything in $\mathcal{B}_{100}(v)$ before time r. More formally, for any update sequence $\{(v_u, p_u)\}_{u=r-T}^{r-1}$ with $v_s = w_1$, define the update sequence $\{(v'_u, p_u)\}_{u=r-T}^{r-1}$ by $v'_u = v_u$ for $u \neq s$ and $v'_s = w_2$. This map is a bijection between the update sequences allowed by the conditions in the numerator of equation (6.27) and the update sequences allowed by the conditions in the denominator of equation (6.27); the existence of this bijection proves

equation (6.27). We similarly observe that if $w_1 \in \mathcal{B}_{100}(v)$ and $w_2 \in G \setminus \mathcal{B}_{108}(v)$, then

$$\frac{\mathbb{P}[v_s = w_1 | \mathcal{S}_t, \mathcal{G}_{v,r}]}{\mathbb{P}[v_s = w_2 | \mathcal{S}_t, \mathcal{G}_{v,r}]} \le 1, \tag{6.28}$$

since certain updates within $\mathcal{B}_{100}(v)$ may be forbidden by the conditioning $\mathcal{G}_{v,r}$. This can be made formal in essentially the same way as the argument for equation (6.27). Inequalities (6.27) and (6.28) imply

$$\begin{aligned}\mathbb{P}\big[v_s \in \mathcal{B}_{100}(v) | \mathcal{S}_t, \mathcal{G}_{v,r}\big] &\le \frac{|\mathcal{B}_{100}(v)|}{n - |\mathcal{B}_{108}(v)|} \\ &\le \frac{(201)^d}{n} + o\big(n^{-1}\big),\end{aligned}$$

where the second line follows from noting that $|\mathcal{B}_\ell(v)| \le (2\ell+1)^d$. Combining this with inequality (6.26),

$$\mathtt{T}_3 = \mathbb{P}\big[\mathcal{A}^{(\text{regular})}_{v,r,C_1\log(n)^4} | \mathcal{S}_t\big] \ge 1 - \frac{(201)^d + (201)^d \log(n)^4}{C_1 \log(n)^4} + o\big(n^{-1}\big).$$

Since $C_1 = 25(201)^d$, for n sufficiently large,

$$\mathtt{T}_3 \ge \frac{23}{25}. \tag{6.29}$$

Next, we bound $\mathtt{T}_2$. For $H \subset \Lambda(L,d)$, let $\phi_H = \{\phi_H(0), \phi_H(1), \ldots\} = \{s \ge t : v_s \in H\}$ be the ordered list of times at which the update vertex v_s falls in H. Then let $\mathcal{H}_{v,r}$ be the σ-algebra generated by the random variables $\{v_s : s \in \phi_{\mathcal{B}_{100}(v)} \cap \{r-T, \ldots, r-1\}\}$. Unlike $\mathcal{G}_{v,r}$, the update times $\phi_{\mathcal{B}_{100}(v)}$ are not included in this σ-algebra, only the update locations v_s.

For notational convenience, let $\mathcal{S}'_t = \{\mathcal{F}_t, \mathcal{A}^{(\text{coll})}_{v,r}, \mathcal{A}^{(\text{sparse})}_{v,r,100}, \mathcal{A}^{(\text{regular})}_{v,r,C_1\log(n)^4}\}$. The gap condition $\mathcal{A}^{(\text{gap})}_{v,r}$ is fulfilled if there are no updates to the region $\mathcal{B}_{100}(v)$ in the time interval $\{r - B_1 n, \ldots, r-1\}$, and so

$$\begin{aligned}\mathtt{T}_2 &= \mathbb{P}\big[\mathcal{A}^{(\text{gap})}_{v,r} | \mathcal{S}'_t\big] \\ &\ge \mathbb{E}\big[\mathbb{P}\big[\psi_{\mathcal{B}_{100}(v)} \cap \{r - B_1 n, \ldots, r-1\} = \varnothing | \mathcal{S}'_t, \mathcal{H}_{v,r}\big] | \mathcal{S}'_t\big].\end{aligned} \tag{6.30}$$

The indices $\phi_{\mathcal{B}_{100}(v)} \cap \{r-T, \ldots, r-1\}$ are, conditional on $\mathcal{H}_{v,r}$, a uniformly-generated size-$|\phi_{\mathcal{B}_{100}(v)} \cap \{r-T, \ldots, r-1\}|$ subset of $\{r-T, \ldots, r-1\}$. Since $|\phi_{\mathcal{B}_{100}(v)} \cap \{r-T, \ldots, r-1\}| \le C_1 \log(n)^4$ by the condition $\mathcal{A}^{(\text{regular})}_{v,r,C_1\log(n)^4}$, we have for n sufficiently large

$$\begin{aligned}&\mathbb{P}\big[\psi_{\mathcal{B}_{100}(v)} \cap \{r - B_1 n, \ldots, r-1\} = \varnothing | \mathcal{S}'_t, \mathcal{H}_{v,r}\big] \\ &\quad = \sum_{0 \le k \le C_1 \log(n)^4} \frac{\binom{T - B_1 n}{k}}{\binom{T}{k}} \mathbb{P}\big[|\phi_{\mathcal{B}_{100}(v)} \cap \{r-T, \ldots, r-1\}| = k | \mathcal{S}'_t, \mathcal{H}_{v,r}\big]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{0\le k\le C_1\log(n)^4} \prod_{i=0}^{B_1 n-1} \frac{T-k-i}{T-i} \\
&\quad \times \mathbb{P}\big[\big|\phi_{\mathcal{B}_{100}(v)} \cap \{r-T,\dots,r-1\}\big| = k \big| \mathcal{S}'_t, \mathcal{H}_{v,r}\big] \\
&\ge \sum_{0\le k\le C_1\log(n)^4} \prod_{i=0}^{B_1 n-1} \left(1-\frac{2k}{T}\right) \\
&\quad \times \mathbb{P}\big[\big|\phi_{\mathcal{B}_{100}(v)} \cap \{r-T,\dots,r-1\}\big| = k \big| \mathcal{S}'_t, \mathcal{H}_{v,r}\big] \\
&\ge \prod_{i=0}^{B_1 n-1} \left(1-\frac{2C_1\log(n)^4}{n\log(n)^4}\right) \ge e^{-4B_1C_1}.
\end{aligned}
$$

Combining this with inequality (6.30) gives

$$
\mathtt{T}_2 \ge e^{-4B_1C_1}. \tag{6.31}
$$

Finally, we bound the term $\mathtt{T}_1$. Let

$$
\mathcal{A}^{(\mathrm{cover})}_{v,r} = \left\{ \mathcal{B}_{100}(v) \subset \bigcup_{r-B_1 n\le s\le r-1} \{v_s\} \right\}
$$

be the event that every element of $\mathcal{B}_{100}(v)$ is updated during the time interval $\{r-B_1n,\dots,r-1\}$. Roughly speaking, for any fixed configuration $X\in\{0,1\}^{\Lambda(L,d)}$, we will denote by $\mathcal{E}^{(\mathrm{coll-up})}_{v,r,X}$ all of the updates "allowed" by $\mathcal{A}^{(\mathrm{sparse})}_{v,r,100}$, $\mathcal{A}^{(\mathrm{gap})}_{v,r}$ and $\mathcal{A}^{(\mathrm{coll})}_{v,r}$. More precisely, $\mathcal{E}^{(\mathrm{coll-up})}_{v,r,X}$ is the set of updates $\{(v_s,p_s)\}_{r-B_1n\le s\le r-1}$ that have the properties:

- If $p_s\le \frac{c}{n}$, then $v_s\notin\mathcal{B}_{100}(v)$.
- If $X_{r-B_1n}=X$ and this KCIP process is updated using the dynamics (1.1) with update variables $\{(v_s,p_s)\}_{r-B_1n\le s\le r-1}$, then $X_{r-1}[v]=0$ and also there exist two neighbours w_1,w_2 of v that are in distinct components of G_{r-1} and satisfy $X_{r-1}[w_1]=X_{r-1}[w_2]=1$.

Similarly, we will denote by $\mathcal{E}^{(\mathrm{cov-up})}_{v,r,X}$ the set of updates "allowed" by $\mathcal{A}^{(\mathrm{sparse})}_{v,r,100}$, $\mathcal{A}^{(\mathrm{gap})}_{v,r}$ and $\mathcal{A}^{(\mathrm{cover})}_{v,r}$. More precisely, this is the set of updates $\{(v_s,p_s)\}_{r-B_1n\le s\le r-1}$ that have the properties:

- If $p_s\le\frac{c}{n}$, then $v_s\notin\mathcal{B}_{100}(v)$.
- $\mathcal{B}_{100}(v)\subset\bigcup_{s=r-B_1n}^{r-1}\{v_s\}$.

We claim that if v,r,X are such that $|\mathcal{E}^{(\mathrm{coll-up})}_{v,r,X}\cap\mathcal{E}^{(\mathrm{cov-up})}_{v,r,X}|\ge 1$, then in fact

$$
\frac{|\mathcal{E}^{(\mathrm{coll-up})}_{v,r,X}\cap\mathcal{E}^{(\mathrm{cov-up})}_{v,r,X}|}{|\mathcal{E}^{(\mathrm{cov-up})}_{v,r,X}|} \ge (2d+1)^{-2d}. \tag{6.32}
$$

To see this, denote by $\{w_i\}_{i=1}^{2d}$ the neighbours of v and by $\{w_{i,j}\}_{j=1}^{2d}$ the neighbours of w_i. Call an element of $\mathcal{E}_{v,r,X}^{(\text{cov}-\text{up})}$ *good*, if for all $1 \le i \le 2d$,

$$\inf\{s : s \ge r - B_1 n, v_s = w_i\} \ge \inf_{1 \le j \le 2d} \inf\{s : s \ge r - B_1 n, v_s = w_{i,j}\}.$$

In other words, a sequence is *good* if vertex w_i is not updated until after all of the vertices $\{w_{i,j}\}_{j=1}^{2d}$ have been updated at least once. A good sequence will not ever remove a particle from a neighbour of v, and the configuration $X_{r-1}|_{\mathcal{B}_{100}(v)}$ resulting from a good sequence will have only singletons. Thus, any good sequence will also be in $\mathcal{E}_{v,r,X}^{(\text{coll}-\text{up})}$ if $\mathcal{E}_{v,r,X}^{(\text{coll}-\text{up})}$ is not empty. Since at least one out of every $(2d+1)^{2d}$ elements of $\mathcal{E}_{v,r,X}^{(\text{cov}-\text{up})}$ is good, this observation implies inequality (6.32).

Inequality (6.32) can be expressed as

$$\begin{aligned}&\mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}, \mathcal{A}_{v,r}^{(\text{cover})}\big]\\ &\qquad \ge (2d+1)^{-(2d+1)}\mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big].\end{aligned}$$

We thus have

$$\begin{aligned}\mathtt{T}_1 &= \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}_{v,r}^{(\text{coll})}, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big]\\ &\ge \mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}_{v,r}^{(\text{coll})}, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}, \mathcal{A}_{v,r}^{(\text{cover})}\big]\\ &\quad \times \mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r}^{(\text{coll})}, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big]\\ &= 1 \times \mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r}^{(\text{coll})}, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big]\\ &= \mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}, \mathcal{A}_{v,r}^{(\text{cover})}\big]\frac{\mathbb{P}[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}]}{\mathbb{P}[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}]}\\ &\ge (2d+1)^{-2d}\mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big]\frac{\mathbb{P}[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}]}{\mathbb{P}[\mathcal{A}_{v,r}^{(\text{coll})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}]}\\ &= (2d+1)^{-2d}\mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big].\end{aligned} \tag{6.33}$$

By a monotonicity argument essentially identical to that used to prove equation (6.27) followed by the standard "coupon-collector" bound, there exists some $0 < B = B(c,d) < \infty$ so that

$$\mathbb{P}\big[\mathcal{A}_{v,r}^{(\text{cover})}|\mathcal{F}_t, \mathcal{A}_{v,r,100}^{(\text{sparse})}, \mathcal{A}_{v,r}^{(\text{gap})}\big] \ge \frac{1}{2} \tag{6.34}$$

for all $B_1 > B$ and all $n > N(B_1)$ sufficiently large. Choosing $B_1 = B + 1$, and combining equality (6.24) with inequalities (6.29), (6.31), (6.33) and (6.34) gives the bound

$$\mathbb{P}\big[\mathcal{M}_{\text{col}}(r) = 1|\mathcal{F}_t, \mathcal{A}_{v,r}^{(\text{coll})}, \mathcal{A}_{v,r,100}^{(\text{sparse})}\big] \ge \frac{23}{50}e^{-4(B+1)C_1}(2d+1)^{-2d} + o\big(n^{-1}\big).$$

Combining this with inequality (6.22), we have

$$\mathbb{P}\big[\mathcal{M}_{\mathrm{col}}(r) = 1 | \mathcal{F}_t, \mathcal{A}_{v,r}^{(\mathrm{coll})}\big] \geq \delta + o(n^{-1})$$

with $\delta = \frac{23}{50} e^{-4(B+1)C_1} (2d+1)^{-2d}$, verifying (6.23). By the observations made in (6.19) and (6.22), the conclusion in (6.18) follows immediately and the proof is completed. □

Recall the definition of the number of collisions $\mathcal{C}_s$ from formula (6.17). We will use Lemma 6.9 to bound the expected change $\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t$ in terms of $\mathcal{C}_{\varepsilon n^3}$:

LEMMA 6.11 (Comparison of number of components to number of collisions). *Fix $\delta = \delta(c,d) > 0$ as given by Lemma 6.9. Then for all $0 < \varepsilon < \varepsilon(c,d)$ sufficiently small,*

$$\mathbb{E}[\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t | \mathcal{F}_t] \leq -\frac{2\delta}{3} \mathbb{E}\big[\big|\{t \leq u \leq t + \varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) \neq 0\}\big| \big| \mathcal{F}_t\big] + \tilde{Y}_t \frac{128 c^2 (d+1)^3 \varepsilon}{3} \big(1 + o(1)\big) + O(1),$$

where the implied constants do not depend on ε or n.

PROOF. Throughout the proof of this lemma, we use "update variables" $\{p_s\}_{s\in\mathbb{N}}$, $\{v_s\}_{s\in\mathbb{N}}$ from formula (1.1). We have

$$\begin{aligned} \mathbb{E}[\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t | \mathcal{F}_t] &= \mathbb{E}\Bigg[\sum_{t \leq u \leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) = 1} (\tilde{Y}_{u+1} - \tilde{Y}_u) \Big| \mathcal{F}_t \Bigg] \\ &\quad + \mathbb{E}\Bigg[\sum_{t \leq u \leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) \neq 1} (\tilde{Y}_{u+1} - \tilde{Y}_u) \Big| \mathcal{F}_t \Bigg] \\ &\equiv \mathrm{T}_1 + \mathrm{T}_2. \end{aligned} \tag{6.35}$$

We first estimate the term T_2. From Definition 6.3, we have that $\mathbb{E}[(\tilde{Y}_{u+1} - \tilde{Y}_u) | p_u > \frac{c}{n}, \mathcal{F}_u] = 0$, and so

$$\mathrm{T}_2 = \mathbb{E}\Bigg[\sum_{t \leq u \leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u) \neq 1} (\tilde{Y}_{u+1} - \tilde{Y}_u) \mathbf{1}_{\mathcal{M}_{\mathrm{col}}(u) \neq 1} \mathbf{1}_{p_u \leq \frac{c}{n}} \Big| \mathcal{F}_t \Bigg]. \tag{6.36}$$

By Lemma 6.7, the corrected component count $\tilde{Y}_u$ cannot increase by more than one when vertices are added:

$$(\tilde{Y}_{u+1} - \tilde{Y}_u) \mathbf{1}_{p_u \leq \frac{c}{n}} \leq 1.$$

Finally, if $|H| \leq 2$, then $\mathcal{N}_H = 1$. In particular, adding a vertex to X_u can only increase $\tilde{Y}_u$ if it connects two components or if it is added to a component that

already has at least two vertices. Since there are at most $4d\delta_u$ vertices adjacent to such a component,

$$\mathbb{E}\big[(\tilde{Y}_{u+1}-\tilde{Y}_u)\mathbf{1}_{\mathcal{M}_{\mathrm{col}}(u)\neq 1}\mathbf{1}_{p_u\leq\frac{c}{n}}|\mathcal{F}_t\big]\leq\frac{4cd}{n^2}\mathbb{E}[\delta_u|\mathcal{F}_t].$$

Combining this bound with inequality (6.36),

$$\mathtt{T}_2\leq\frac{4cd}{n^2}\mathbb{E}\bigg[\sum_{u=t}^{t+\varepsilon n^3}\delta_u\Big|\mathcal{F}_t\bigg]. \tag{6.37}$$

By inequality (6.15) and the calculation in inequality (6.16), we have

$$\begin{aligned}\frac{4cd}{n^2}\mathbb{E}\bigg[\sum_{u=t}^{t+\varepsilon n^3}\delta_u|\mathcal{F}_t\bigg]&\leq\tilde{Y}_t\frac{128c^2(d+1)^3}{n^3}\bigg(8n\bigg(1-\frac{1}{8n}\bigg)^{\varepsilon n^3}+\varepsilon n^3\bigg)\\&\quad+\frac{32cd}{n}\delta_t+\frac{4}{3}\bigg(1-\frac{1}{8n}\bigg)^{\varepsilon n^3}\delta_t.\end{aligned}$$

Combining this with inequality (6.37), this implies

$$\begin{aligned}\mathtt{T}_2&\leq\tilde{Y}_t\frac{128c^2(d+1)^3}{3n^3}\bigg(8n\bigg(1-\frac{1}{8n}\bigg)^{\varepsilon n^3}+\varepsilon n^3\bigg)\\&\quad+\frac{32cd}{n}\delta_t+\bigg(1-\frac{1}{8n}\bigg)^{\varepsilon n^3}\delta_t\\&\leq\tilde{Y}_t\frac{128c^2(d+1)^3\varepsilon}{3}\big(1+o(1)\big)+O(1).\end{aligned} \tag{6.38}$$

Next, we turn to bounding $\mathtt{T}_1$. By Lemma 6.6, it follows that

$$\begin{aligned}\mathtt{T}_1&=\mathbb{E}\bigg[\sum_{t\leq u\leq t+\varepsilon n^3:\mathcal{M}_{\mathrm{col}}(u)=1}(\tilde{Y}_{u+1}-\tilde{Y}_u)\Big|\mathcal{F}_t\bigg]\\&\leq-\frac{2}{3}\mathbb{E}\big[|\{t\leq u\leq t+\varepsilon n^3:\mathcal{M}_{\mathrm{col}}(u)=1\}||\mathcal{F}_t\big]+0.\end{aligned}$$

By Lemma 6.9, this implies

$$\begin{aligned}\mathtt{T}_1&\leq-\frac{2\delta}{3}\mathbb{E}\big[|\{t+n\log(n)^4\leq u\leq t+\varepsilon n^3:\mathcal{M}_{\mathrm{col}}(u)\neq 0\}||\mathcal{F}_t\big]\\&\quad+o\big(n^{-1}\big).\end{aligned} \tag{6.39}$$

Thus,

$$\begin{aligned}&\mathbb{E}\big[|\mathcal{T}_{\mathrm{col}}(\varepsilon n^3)\cap\{t,t+1,\ldots,t+n\log(n)^4\}||\mathcal{F}_t\big]\\&\quad\leq\mathbb{E}\big[|\{t\leq u\leq t+n\log(n)^4:V_{u+1}>V_u\}||\mathcal{F}_t\big]\end{aligned}$$

$$\leq \frac{2cd}{n^2}\mathbb{E}\Bigg[\sum_{u=t}^{t+n\log(n)^4} V_u|\mathcal{F}_t\Bigg]$$

$$\leq \frac{4cd(d+1)}{n^2}\mathbb{E}\Bigg[\sum_{u=t}^{t+n\log(n)^4} \tilde{Y}_u\Big|\mathcal{F}_t\Bigg] \tag{6.40}$$

$$\leq \frac{4cd(d+1)}{n^2}\sum_{v=0}^{n\log(n)^4}\Bigg(\tilde{Y}_t\Bigg(1+\frac{96c^2(d+1)^3}{n^3}\Bigg(8n\Bigg(1-\frac{1}{8n}\Bigg)^v+v\Bigg)\Bigg)$$
$$+\frac{24cd}{n}\delta_t+\Bigg(1-\frac{1}{8n}\Bigg)^v\delta_t\Bigg)$$

$$\leq \tilde{Y}_t\big(1+o(1)\big)\frac{\log(n)^4}{n}+O(1),$$

where the third inequality is due to inequality (6.14) and the fourth inequality is due to Lemma 6.8. Combining inequalities (6.39) and (6.40), we have

$$\mathrm{T}_1 \leq -\frac{2\delta}{3}\mathbb{E}\big[\big|\{t\leq u\leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u)\neq 0\}\big||\mathcal{F}_t\big]$$
$$+\frac{2\delta}{3}\tilde{Y}_t\big(1+o(1)\big)\frac{\log(n)^4}{n}+O(1).$$

Combining this inequality with inequalities (6.38) and (6.35), we have

$$\mathbb{E}[\tilde{Y}_{t+\varepsilon n^3}-\tilde{Y}_t|\mathcal{F}_t]$$
$$\leq -\frac{2\delta}{3}\mathbb{E}\big[\big|\{t\leq u\leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u)\neq 0\}\big||\mathcal{F}_t\big]$$
$$+\frac{2\delta}{3}\tilde{Y}_t\big(1+o(1)\big)\frac{\log(n)^4}{n}+\tilde{Y}_t\frac{128c^2(d+1)^3\varepsilon}{3}\big(1+o(1)\big)+O(1)$$
$$=-\frac{2\delta}{3}\mathbb{E}\big[\big|\{t\leq u\leq t+\varepsilon n^3 : \mathcal{M}_{\mathrm{col}}(u)\neq 0\}\big||\mathcal{F}_t\big]$$
$$+\tilde{Y}_t\frac{128c^2(d+1)^3\varepsilon}{3}\big(1+o(1)\big)+O(1),$$

and the proof is completed. □

6.4. *Colored constrained Ising process.* Our next goal is to prove a lower bound on the expected number of collisions. To this end, we define a "colored" version of the KCIP on a general finite graph G, which allows us to make rigorous the notion of a single particle moving and branching over time.

DEFINITION 6.12 (Colored constrained Ising process). Fix $t\in\mathbb{N}$, $1\leq k\leq n$ and $x\in\{0,1\}^G$ so that the subgraph of G induced by the vertices $\{v\in G : x[v]=$

1} has exactly k connected components. We define a Markov chain $\{\widehat{X}_s\}_{s\geq t}$ on the state space $\{0,1,\ldots,k\}^n$ that is closely coupled to the KCIP $\{X_s\}_{s\geq t}$ started at $X_t = x$; in fact, we will have $X_s[v] = \mathbf{1}_{\widehat{X}_s[v]\neq 0}$ for all $v \in G$ and $s \geq t$. We begin by setting the initial condition $\widehat{X}_t$. For a fixed $v \in G$, if $X_t[v] = 0$, we also set $\widehat{X}_t[v] = 0$. Fix an ordering of the k connected components $c_t[1],\ldots,c_t[k]$ of G_t, and set $\widehat{X}_t[v] = i$ for all $v \in c_t[i]$ and $\widehat{X}_t[v] = 0$ for all $v \notin \bigcup_{i=1}^k c_t[i]$. Note that this arbitrary ordering and labelling of the components is done once, at time t. We do not reorder components at times $s > t$, and we will always have $\widehat{X}_s[v]$ in the set of labels $\{0,1,\ldots,k\}$, even if the number of components at time s is not equal to k. Indeed, it will turn out that with probability one there exists a (random) index $i \in \{1,2,\ldots,k\}$ and time $S \geq t$ so that $\widehat{X}_s[v] \in \{0,i\}$ for all $s > S$.

To evolve $\widehat{X}_s$, recall that $\{X_s\}_{s\geq t}$ evolves by selecting at every time s a vertex v_s to update, and sometimes changing the label of that vertex. Whenever the labelling of a vertex v is changed from 1 to 0 in X_s, the labelling of v should also be changed to 0 in $\widehat{X}_s$. Whenever the labelling of a vertex v is changed from 0 to 1 at time s in X_s, choose a vertex $u_s \sim \mathrm{Unif}\{w : (w,v) \in E, X_s[w] = 1\}$ uniformly at random from the neighbours of v that have a nonzero label in $\widehat{X}_s$, set $\widehat{X}_{s+1}[v] = \widehat{X}_s[u_s]$, and then set $\widehat{X}_{s+1}[w] = \widehat{X}_s[u_s]$ for all $w \in \mathrm{Comp}_{s+1}(u_s)$. All other labels of $\widehat{X}_{s+1}$ should be the same as that of $\widehat{X}_s$. Since entire components can "flip" colors, the process $\{\widehat{X}_s\}_{s\geq t}$ may have many labels change at once.

Recall that a *collision* occurs at time r and vertex u in the original KCIP $\{X_t\}_{t\geq 0}$ if $v_r = u$ and $Y_{r+1} < Y_r$. Analogously, we say that color i is involved in a collision at time r and vertex u if a collision occurs at time r and vertex u, and furthermore there exists $(u,v) \in E$ s.t. $\hat{X}_r[v] = i$. Note that several colors can be involved in one collision.

We now give some definitions related to the colored KCIP. For all $s \geq t$ define $\mathrm{Comp}_s^{(i)}$ to be the connected components of G_s containing only vertices u satisfying $\widehat{X}_s[u] = i$. Define the *number of vertices with color i* by

$$V_s^{(i)} = \sum_{u\in\Lambda(L,d)} \mathbf{1}_{\widehat{X}_s[u]=i}$$

and the associated *interference time* by

$$\zeta_{\mathrm{int}}^{(i)} = \inf\bigl\{s > t : \bigl\{\bigl|V_s^{(i)} - V_{s-1}^{(i)}\bigr| > 1\bigr\} \cup \bigl\{\bigl||\mathrm{Comp}_s^{(i)}| - |\mathrm{Comp}_{s-1}^{(i)}|\bigr| > 0\bigr\}\bigr\}.$$

We also provide a generalization of the definition of the *triple* time in formula (4.4):

$$\zeta_{\mathrm{triple}}^{(i)} = \inf\bigl\{s > t : V_s^{(i)} \geq 3\bigr\}. \tag{6.41}$$

This allows us to state the following corollary of Lemma 4.1.

COROLLARY 5. *Fix $i, t \in \mathbb{N}$ and $\varepsilon > 0$ and assume that $V_t^{(i)} = 1$. Then*

$$\mathbb{P}\left[\zeta_{\text{triple}}^{(i)} - t < \min\left(\varepsilon \frac{n^3}{6c^2 d(2d-1)}, \zeta_{\text{int}}^{(i)}\right)\right] = O(\varepsilon). \tag{6.42}$$

PROOF. We reduce this to the case of Lemma 4.1. Let $v \in \Lambda(L, d)$ be the unique vertex with $\widehat{X}_t[v] = i$. If $\sum_{u \in \Lambda(L,d)} X_t[u] = 1$, we are exactly in the case of Lemma 4.1. If $\sum_{u \in \Lambda(L,d)} X_t[u] > 1$, we will couple $\{X_s\}_{s \geq t}$ to a second KCIP $\{X'_s\}_{s \geq t}$ as follows. We set the initial condition $X'_t[u] = \mathbf{1}_{u=v}$. We then couple $\{X'_s\}_{s \geq t}$ to $\{X_s\}_{s \geq t}$ by using the same update sequence p_s, v_s, as described in representation (1.1).

Observe that, if there exists some time $s > t$ at which $|\{v : X'_s[v] = i\}| \geq 3$, it must be the case that either $\zeta_{\text{triple}}^{(i)} \leq s$ or the color i is involved in a collision at some time $t < r \leq s$. Thus, for any constant $0 < C < \infty$, if $\zeta_{\text{triple}}^{(i)} < \min(C, \zeta_{\text{int}}^{(i)})$ holds for the process $\{X_s\}_{s \geq t}$, it must also hold for the process $\{X'_s\}_{s \geq t}$. Thus, for the purposes of proving inequality (6.42), it is enough to prove it in the case that $X_t[u] = \mathbf{1}_{u=v}$; but in that case, the conclusion follows by Lemma 4.1. □

We next recall the coalescent process [8, 17].

DEFINITION 6.13 (Coalescent process). Fix a graph G and parameters $k \in \mathbb{N}$, $q \in [0, \frac{1}{k}]$. A *coalescent process on graph G with k initial particles and moving rate q* is a Markov chain $\{Z_s\}_{s \in \mathbb{N}}$ on G^k. Let $O_s = \{v \in G : \exists 1 \leq i \leq k \text{ such that } Z_s[i] = v\}$ be the *occupied sites* of Z_s. To evolve Z_s, we first choose $u_s \sim \text{Unif}([0, 1])$, $v_s \sim \text{Unif}([O_s])$ and $u_w \sim \text{Unif}([\mathcal{B}_1(v_s) \setminus \{v_s\}])$ and set $I_s = \{i : Z_s[i] = v_s\}$. If $u_s \leq q|O_s|$, then set $Z_{s+1}[i] = u_w$ for all $i \in I_s$ and set $Z_{s+1}[j] = Z_s[j]$ for all $j \notin I_s$; otherwise, set $Z_{s+1}[j] = Z_s[j]$ for all j.

REMARK 6.14. The coalescent process has many other names and descriptions. The construction of the coalescent process as the "dual" to the voter process is well known (see [17]). Informally, we can view the coalescent process with k initial particles as k random walkers that take turns making independent simple random walk steps until a collision occurs, at which point the colliding particles "merge" into a single particle. After this collision, the coalescent process resumes with $k - 1$ particles.

6.5. *Coupling KCIP with coalescent process.* Next, we couple a colored KCIP $\{\widehat{X}_s\}_{s \geq t}$ with a collection of lazy random walks $\{Q_s[i]\}$ and a coalescent process $\{R_s\}_{s \geq t + n^{2.5}}$ to obtain the following lower bound on the number of collisions $\mathcal{C}_s$ [see formula (6.17)], which is the only result from this section that we will need.

LEMMA 6.15. *With notation as above, there exist constants* $\kappa = \kappa(\varepsilon, c, d) > 0$ *and* $C = C(\varepsilon, c, d) < \infty$ *that do not depend on* n *so that, for any initial configuration* $X_t \in \Omega$,

$$\mathbb{E}[\mathcal{C}_{t+\varepsilon n^3} | \mathcal{F}_t] \geq \kappa Y_t - C.$$

The idea behind the proof of Lemma 6.15 is to show the following four bounds:

1. With high probability, we have $Q_s[i] \in \{v : \widehat{X}_s[v] = i\}$ and $Q_s[i] \in \{v : R_s[v] = 1\}$ for "many" $i \in \{1, 2, \ldots, Y_t\}$ over a "large" interval $s \in \{t + n^{2.5}, t + n^{2.5} + 1, \ldots, T(i)\}$ (see Lemma 6.17).
2. The expected number of "near-collisions" in the KCIP is at least some fixed fraction of the expected number of collisions involving the coalescent process $\{R_s\}_{s \geq t+n^{2.5}}$ (see Proposition 6.19). This occurs because, per **(1)** above, the particles in the coalescent process are often "covered" by the particles in the KCIP. See Figure 6 for an illustration of the relationship between $\widehat{X}_s$ and R_s.
3. Theorem 5 of [9] implies that the expected number of collisions in the coalescent process between time t and $t + \varepsilon n^3$ is "almost" as large as Y_t.
4. The expected number of collisions in the KCIP is at least some fixed fraction of the expected number of near-collisions (see Lemma 6.20).

Before beginning the proof, we construct our couplings and give some relevant definitions. Almost all of these constructions are used only in this section.

6.5.1. *Coupling KCIP to a collection of (lazy) simple random walks.* Assume $n > c$ and fix an index $1 \leq i \leq Y_t$. Let

$$\tau_{\text{start}}^{(i)} = \inf\{s \geq t : \forall w_1, w_2 \in \Lambda(L, d), \{\widehat{X}_s[w_1] = \widehat{X}_s[w_2] = i\} \implies \{(w_1, w_2) \notin E\}\}$$

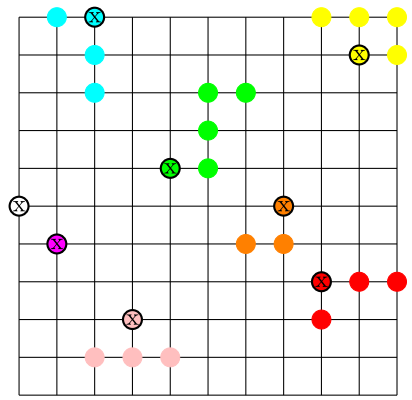

FIG. 6. *Embedded Coalescent Process*: *The colored KCIP with* $k = 7$ *remaining colors is indicated by the colored circles*; *the coupled coalescent process is indicated by the X's. As is typical early in the coupling, most X's lie on vertices covered by the colored KCIP.*

be the first time that no two particles colored i are adjacent. We now construct a coupling of processes $\{Q_s[i]\}_{1\le i\le Y_t, s\ge\tau^{(i)}_{\text{start}}}$ on $\Lambda(L,d)$ to $\{\widehat{X}_s\}_{s\ge t}$ as follows. These processes will each evolve as $(1-\frac{cd}{n^2})$-lazy simple random walks on $\Lambda(L,d)$.

Define the *lifetime* $\ell_s[v]$ of a particle at v with $\widehat{X}_s[v]=i$ to be

$$\ell_s[v]=\sup\{T\ge s:\forall s<u\le T, v_u\ne v\};$$

the right-hand side is positive with probability 1. Fix $1\le i\le Y_t$. If the set $\{v:\widehat{X}_{\tau^{(i)}_{\text{start}}}[v]=i\}$ is nonempty, choose a vertex v uniformly at random from that set and let $Q_{\tau^{(i)}_{\text{start}}}[i]=v$. If that set is empty, choose a vertex v uniformly at random from $\Lambda(L,d)$ and let $Q_{\tau^{(i)}_{\text{start}}}[i]=v$. For $s\ge\tau^{(i)}_{\text{start}}$, we evolve $Q_s[i]$ by always setting

$$Q_{s+1}[i]=\operatorname*{argmax}_{v\in v_s, Q_s[i]}\ell_s[v]. \tag{6.43}$$

For our purposes, we are not interested in the process $Q_s[i]$ for all $s\ge\tau^{(i)}_{\text{start}}$ but only until a "decoupling time." To this end, define the random time

$$\tau^{(i)}_{\text{triple}}=\inf\big\{s>\tau^{(i)}_{\text{start}}:\text{the component of } G_s \text{ containing } Q_s[i]\ \text{has at least three elements.}\big\}.$$

Next, define the decoupling time of color i to be the minimum of the first time that the particle at $Q_s[i]$ does not have color i and the first time that there is a size-three component of color i:

$$\tau^{(i)}_{\text{decoupling}}=\min\big(\tau^{(i)}_{\text{triple}},\inf\{s\ge\tau^{(i)}_{\text{start}}:\widehat{X}_s[Q_s[i]]\ne i\}\big). \tag{6.44}$$

LEMMA 6.16. *For any fixed $1\le i\le Y_t$, the process $\{Q_s[i]\}_{s\ge\tau^{(i)}_{\text{start}}}$ has the distribution of a $(1-\frac{cd}{n^2})$-lazy random walk on $\Lambda(L,d)$.*

PROOF. Let $\{(v_s,p_s)\}_{s\ge t}$ be the update variables used for the KCIP, as in equation (1.1). It is a direct computation that for $s\ge\tau^{(i)}_{\text{start}}$ and vertices $u\ne v$ satisfying $|u-v|=1$,

$$\begin{aligned}
&\mathbb{P}\big[Q_{s+1}[i]=v|Q_s[i]=u\big]\\
&\quad=\mathbb{P}\bigg[\{v_s=v\}\cap\bigg\{p_s\le\frac{c}{n}\bigg\}\cap\big\{\ell_s(v)>\ell_s(u)\big\}\bigg]\\
&\quad=\mathbb{P}[v_s=v]\mathbb{P}\bigg[p_s\le\frac{c}{n}\bigg]\mathbb{P}\bigg[\ell_s(v)>\ell_s(u)|v_s=v,\,p_s\le\frac{c}{n}\bigg]\\
&\quad=\frac{1}{n}\frac{c}{n}\frac{1}{2}=\frac{c}{2n^2}.
\end{aligned}$$

It is immediate that $\mathbb{P}[Q_{s+1}[i]=u|Q_s[i]=u]=1-\frac{cd}{n^2}$. These two inequalities prove the result. □

6.5.2. *Near collisions*. Define the "near-collision" time associated with color i to be the first time at which a vertex that has color i is exactly distance 2 from another particle in the KCIP:

$$\tau_{\text{near}}^{(i)} = \inf\{s \geq \tau_{\text{start}}^{(i)} : \exists u, v \in \Lambda(L,d) \text{ s.t. } |u-v| = 2 \quad \text{and} \quad X_s[u] = 1 \quad \text{and} \quad \hat{X}_s[v] = i\}.$$

Similarly, say that color $1 \leq i \leq Y_t$ has a *near-collision* between times t_1 and t_2 if there exists a time $t_1 \leq s \leq t_2$ and a pair of vertices u, v so that $\widehat{X}_s[u] = i$ and $X_s[v] = 1$, so that u, v are not in the same component of G_s, and so that $|u - v| = 2$.

Define

$$\mathcal{Q}_s = \{t \leq r \leq s : \exists u, v \in \Lambda(L,d) \text{ s.t.} |u-v| = 2 \quad \text{and} \quad X_r[u] = X_r[v] = 1 \quad \text{and} \quad u \notin \text{Comp}_r(v)\} \tag{6.45}$$

to be the set of times at which a near-collision occurs and let

$$\mathcal{N}_s^{\text{near}} = |\mathcal{Q}_s| \tag{6.46}$$

be the total number of near-collisions between times t and s.

We have the following lemma.

LEMMA 6.17. *For* $1 \leq i \leq Y_t$,

$$\tau_{\text{decoupling}}^{(i)} \geq \min(\tau_{\text{triple}}^{(i)}, \tau_{\text{near}}^{(i)}). \tag{6.47}$$

PROOF. To prove inequality (6.47), observe that it can only be violated if the particle $Q_s[i]$ becomes "uncovered" before $\tau_{\text{triple}}^{(i)}$ or $\tau_{\text{near}}^{(i)}$:

$$\tau_{\text{uncovered}}^{(i)} \equiv \inf\{s \geq \tau_{\text{start}}^{(i)} : \widehat{X}_s[Q_s[i]] \neq i\} < \min(\tau_{\text{triple}}^{(i)}, \tau_{\text{near}}^{(i)}). \tag{6.48}$$

By the construction (6.43), we can only have $\tau_{\text{uncovered}}^{(i)} < \tau_{\text{triple}}^{(i)}$ if there is a collision involving color i at time $\tau_{\text{uncovered}}^{(i)}$. However, any collision involving color i must be preceded by a near-collision—more precisely, if there is a collision involving color i at time $\tau_{\text{uncovered}}^{(i)}$, there must be some time $r < \tau_{\text{uncovered}}$ at which there exist $u, v \in \Lambda(L,d)$ such that

$$|u-v| = 2 \quad \text{and} \quad X_s[u] = 1 \quad \text{and} \quad \hat{X}_s[v] = i.$$

But this means exactly that

$$\{\tau_{\text{uncovered}}^{(i)} < \tau_{\text{triple}}^{(i)}\} \Rightarrow \{\text{there is a collision involving color } i \text{ at time } \tau_{\text{uncovered}}\} \Rightarrow \{\tau_{\text{uncovered}} > \tau_{\text{near}}^{(i)}\}.$$

Combining this implication with inequality (6.48) completes the proof of inequality (6.47). □

6.5.3. *Coupling Q_s to a coalescent process.* Having defined the Q_s processes, we couple them to a coalescent process $\{R_s\}_{s\geq t+n^{2.5}}$ with Y_t initial particles. We begin by setting the initial conditions $R_{t+n^{2.5}}$. If $\tau^{(i)}_{\text{start}} \leq n^{2.5}$, we set $R_{t+n^{2.5}}[i] = Q_{t+n^{2.5}}[i]$. Otherwise, choose $R_{t+n^{2.5}}[i]$ uniformly at random from among all vertices in $\Lambda(L,d) \setminus \bigcup_i \{Q_{t+n^{2.5}}[i]\}$.

We now describe the evolution of $\{R_s\}_{s\geq t+n^{2.5}}$. Let $\mathcal{D} \subset \{1,2,\ldots,Y_t\} \times \{t, t+1, \ldots\}$ denote the pairs of indices (i,s) that satisfy

$$\tau^{(i)}_{\text{start}} \leq t + n^{2.5} \leq s \leq \min\bigl(\tau^{(i)}_{\text{triple}}, \tau^{(i)}_{\text{near}}\bigr). \tag{6.49}$$

For $(i,s) \in \mathcal{D}$, we set $R_s[i] = Q_s[i]$. For $(i,s) \notin \mathcal{D}$, we choose $R_{s+1}[i]$ conditional on R_s and $\{R_{s+1}[j]\}_{(j,s)\in\mathcal{D}}$ and independently of all other random variables being discussed.

This defines the evolution of $R_s[i]$ for $s, i \in \mathcal{D}$. Note that the times $\tau^{(i)}_{\text{start}}$, $\tau^{(i)}_{\text{triple}}$, and $\tau^{(i)}_{\text{near}}$ that determine the boundaries of $\mathcal{D}$ are all stopping times with respect to the joint evolution of $\{R_s, \widehat{X}_s\}$. Thus, by standard arguments, it is possible to extend our construction of the coalescent process $\{R_s[i]\}$ from pairs $s, i \in \mathcal{D}$ to all pairs $1 \leq i \leq Y_t$, $t + n^{2.5} \leq s$.

REMARK 6.18. The coupling given in Section 6.5.1 has the critical property that $\widehat{X}_s[R_s[i]] = i$ for any $1 \leq i \leq Y_t$ and $t + n^{2.5} \leq s < \tau^{(i)}_{\text{decoupling}}$.

We make some further observations about the above construction. Say that color i has *coalesced by time* t_1 if $R_s[i] = R_s[j]$ for some $j \neq i$ and $s \leq t_1$. Define the events:

1. $\mathcal{A}^{(i)}_{1,u}$: $\tau^{(i)}_{\text{start}} > t + u$ and color i has no near-collisions between time t and $t+u$.
2. $\mathcal{A}^{(i)}_2$: $\tau^{(i)}_{\text{start}} \leq t + n^{2.5}$ but color i has not coalesced by time $t + \varepsilon n^3$.
3. $\mathcal{A}^{(i)}_3$: $\tau^{(i)}_{\text{triple}} < \min(\tau^{(i)}_{\text{near}}, \varepsilon n^3)$.
4. $\mathcal{A}^{(i)}_4$: Denote by $J \subset \{1,2,\ldots,Y_t\}$ the set of all colors j s.t. for some $t \leq s \leq \varepsilon n^3$, we have $R_s[i] = R_s[j]$. $\mathcal{A}^{(i)}_4$ is the event that $\tau^{(i)}_{\text{start}} \leq t + n^{2.5}$ and color i has coalesced, but for all $j \in J$, we also have $\tau^{(j)}_{\text{decoupling}} < \tau^{(i)}_{\text{near}}$.

We have the following.

PROPOSITION 6.19. *The color i has a near-collision between times t and $t + \varepsilon n^3$ unless at least one of the events $\mathcal{A}^{(i)}_1 \equiv \mathcal{A}^{(i)}_{1,n^{2.5}}$, $\mathcal{A}^{(i)}_2$, $\mathcal{A}^{(i)}_3$ or $\mathcal{A}^{(i)}_4$ occurs.*

PROOF. Fix $1 \leq i \leq Y_t$ and denote by $\mathcal{B}^{(i)}$ the event that color i has a near-collision between times t and $t + \varepsilon n^3$. We assume that none of the events $\{\mathcal{A}^{(i)}_j\}_{j=1}^4$ nor $\mathcal{B}^{(i)}$ occur, and will derive a contradiction.

Since $\mathcal{A}_1^{(i)}$ does not occur, we have either $\widehat{X}_s[R_s[i]] = i$ for all

$$t + n^{2.5} \le s < \tau^{(i)}_{\text{decoupling}} \tag{6.50}$$

or that color i has a near-collision by time $n^{2.5}$. Since we have assumed that $\mathcal{B}^{(i)}$ does not occur, this implies by inequality (6.47) that $\widehat{X}_s[R_s[i]] = i$ for all

$$t + n^{2.5} \le s < \min\bigl(\tau^{(i)}_{\text{triple}}, \tau^{(i)}_{\text{near}}\bigr). \tag{6.51}$$

Since $\mathcal{B}^{(i)}$ does not occur, we have $\tau^{(i)}_{\text{near}} \ge \varepsilon n^3$, and so by inequality (6.51) we have $\widehat{X}_s[R_s[i]] = i$ for all

$$t + n^{2.5} \le s < \min\bigl(\tau^{(i)}_{\text{triple}}, \varepsilon n^3\bigr). \tag{6.52}$$

Since neither $\mathcal{A}_3^{(i)}$ nor $\mathcal{B}^{(i)}$ occur, we have that $\tau^{(i)}_{\text{triple}} \ge \varepsilon n^3$. Thus, by inequality (6.52) we have that $\widehat{X}_s[R_s[i]] = i$ for all

$$t + n^{2.5} \le s < \varepsilon n^3. \tag{6.53}$$

However, since $\mathcal{A}_2^{(i)}$ does not occur, we have that color i has coalesced at some time $t + n^{2.5} \le s \le \varepsilon n^3$. Since particles must be nearby before they can coalesce, this means that there is some set $J \subset \{1, 2, \ldots, Y_t\}$, some times $\{r_j\}_{j\in J}$ satisfying $t + n^{2.5} \le r_j < s \le \varepsilon n^3$, and some vertex $\{v_j\}_{j\in J} \subset \Lambda(L, d)$ so that

$$R_{r_j}[j] = v_j, \qquad \bigl|v_j - R_{r_j}[i]\bigr| = 2.$$

Since $\mathcal{B}^{(i)}$ does not occur, this implies that $\tau^{(j)}_{\text{decoupling}} \ge r_j$ for all $j \in J$. However, this contradicts our assumption that $\mathcal{A}_4^{(i)}$ does not occur. This completes the proof of the proposition. □

We must relate the number of near-collisions $\mathcal{N}^{\text{near}}_{t+\varepsilon n^3}$ to the number of collisions $\mathcal{C}_{t+\varepsilon n^3}$.

PROPOSITION 6.20. *Fix $\varepsilon > 0$. There exists $\kappa = \kappa(c, d) > 0$ so that for all n sufficiently large,*

$$\mathbb{E}[\mathcal{C}_{t+\varepsilon n^3+4n^{2.5}}] \ge \kappa \mathbb{E}\bigl[\mathcal{N}^{\text{near}}_{t+\varepsilon n^3}\bigr].$$

PROOF. It is enough to check that, once two components C_1, C_2 of G_s have vertices $w_1 \in C_1, w_2 \in C_2$ with $|w_1 - w_2| = 2$, there is a positive probability that all three of the following events occur:

- The particles at w_i survive longer than any of their neighbours.
- A particle is added to the common neighbour v of w_1, w_2 before any other particle is added to the set $\mathcal{B}_1(w_1) \cup \mathcal{B}_1(w_2)$.
- Some particle is added to the set $\mathcal{B}_1(w_1) \cup \mathcal{B}_1(w_2)$ before time $t + 4n^{2.5}$.

We note that the first two events are purely local [i.e., they depend only on the order that events occur within some bounded region of the graph $\Lambda(L,d)$ that does not depend on n], and the last occurs with probability at least $1-(1-\frac{c}{n^2})^{n^{2.5}} \approx 1-e^{-c\sqrt{n}}$. For this reason, it seems clear that this conclusion should hold with some constant; as such we give only a brief formal proof.

Fix a time $t \le s \le t+\varepsilon n^3$ at which a near-collision occurs. Then there exists a triplet of vertices $u, v, w \in G$ so that $X_s[u] = X_s[w] = 1$, u, w are in different components of G_s, and u, w are both adjacent to v. Let

$$\phi_u = \inf\Big\{ r \ge s : \sum_{x \in \mathcal{B}_1(u)} X_r[x] = 1 \text{ or } X_r[u] = 0 \Big\}$$

be the first time after s that u is an isolated vertex or is empty, and let

$$\phi'_u = \inf\{ r \ge \phi_u : X_r|_{\mathcal{B}_1(u)}[x] \ne \mathbf{1}_{x=u} \}$$

be the first time after ϕ_u that it is not an isolated vertex. Define ϕ_w, ϕ'_w analogously. Finally, let $\phi_{\text{bad}} = \inf\{t > s : X_t[u]X_t[w] = 0\}$ and let $\eta_v = \inf\{t > s : X_t[v] = 1\}$. Let $\mathcal{E}_v = \{\eta_v < \min(s+4n^{2.5}, \phi_{\text{bad}})\}$. We have:

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_v] &\ge \mathbb{P}\big[\eta_v < \min(s+4n^{2.5}, \phi_{\text{bad}}) | \phi_u, \phi_w < \min(s+n^{2.5}, \phi_{\text{bad}})\big] \\
&\quad \times \mathbb{P}\big[\phi_u, \phi_w < \min(s+n^{2.5}, \phi_{\text{bad}})\big] \\
&\ge \mathbb{P}\big[\eta_v \le \phi'_u, \phi'_v < \min(s+4n^{2.5}, \phi_{\text{bad}}) | \phi_u, \phi_w < \min(s+n^{2.5}, \phi_{\text{bad}})\big] \\
&\quad \times \mathbb{P}\big[\phi_u, \phi_w < \min(s+n^{2.5}, \phi_{\text{bad}})\big] \qquad (6.54) \\
&\ge \Big(\frac{1}{4d^2} - \mathbb{P}\big[\max(\phi'_u - \phi_u, \phi'_w - \phi_w) > 3n^{2.5}\big]\Big) \\
&\quad \times \Big(\frac{1}{4d^2} - \mathbb{P}\big[\max(\phi_u, \phi_w) - s > n^{2.5}\big]\Big) \\
&\ge \Big(\frac{1}{4d^2} - e^{c\sqrt{n}}\Big)\Big(\frac{1}{4d^2} - 4de^{-(1-c)n\sqrt{n}}\Big).
\end{aligned}$$

We now apply this calculation. Recall from equation (6.45) that $\mathcal{Q}_{\varepsilon n^3}$ is the set of near-collision times. If $s \in \mathcal{Q}_{\varepsilon n^3}$, there exist two components C_1, C_2 of G_s and vertices $u \in C_1$, $w \in C_2$, $v \notin C_1 \cup C_2$ with $v = \mathcal{B}_1(u) \cap \mathcal{B}_1(w)$; for $s \in \mathcal{Q}_{\varepsilon n^3}$, define $f(s) = v$. When $\mathcal{E}_{f(s)}$ holds, define $g(s) = \inf\{r > s : v_r = f(s), p_r \le \frac{c}{n}\}$ to be the time at which the near-collision started at time s is completed to a collision. Noting from the definition of $\mathcal{E}_v$ that $|\{s : g(s) = r\}| \le 2d-1$ for any $r \in \mathbb{N}$, inequality (6.54) then gives

$$\begin{aligned}
\mathbb{E}[\mathcal{C}_{t+\varepsilon n^3+4n^{2.5}}] &\ge \mathbb{E}\Big[\frac{1}{2d-1} \sum_{s \in \mathcal{Q}_{\varepsilon n^3}} \mathbf{1}_{\mathcal{E}_{f(s)}}\Big] \\
&\ge \frac{1}{2d-1}\Big(\frac{1}{4d^2} - e^{c\sqrt{n}}\Big)\Big(\frac{1}{4d^2} - 4de^{-(1-c)n\sqrt{n}}\Big)\mathbb{E}[\mathcal{N}_{\varepsilon n^3}].
\end{aligned}$$

This completes the proof. □

We now prove the main result of this section.

PROOF OF LEMMA 6.15. We will bound the probabilities of the four events $\{\mathcal{A}_j^{(i)}\}_{j=1}^4$ from the statement of Proposition 6.19. Noting that color i has a near-collision if there are ever two components of color i (since the two components must be at distance exactly 2 when they are first separated), we have by essentially the same calculation as inequality (6.5):

$$\begin{aligned}\mathbb{E}\big[V_{t+s+1}^{(i)}\mathbf{1}_{\mathcal{A}_{1,s+1}^{(i)}}|\mathcal{F}_{t+s}\big] &\le V_{t+s+1}^{(i)} - \frac{V_{t+s}^{(i)}}{n}\left(1-\frac{c}{n}\right) + \frac{2dV_{t+s}^{(i)}}{n}\frac{c}{n} \\ &\le V_{t+s}^{(i)}\left(1-\frac{1}{n}\left(1-\frac{c(2d+1)}{n}\right)\right).\end{aligned}$$

Thus, for $n > 2c(2d+1)$, we iterate and find

$$\mathbb{E}\big[V_{t+n^{2.5}}^{(i)}\mathbf{1}_{\mathcal{A}_{1,n^{2.5}}^{(i)}}|\mathcal{F}_t\big] \le ne^{-n}.$$

Since $V_{t+s}^{(i)}$ is at least 1 for all $t+s \le \tau_{\text{near}}^{(i)}$, applying Markov's inequality gives

$$\mathbb{P}\big[\mathcal{A}_{1,n^{2.5}}^{(i)}|\mathcal{F}_t\big] \le \mathbb{E}\big[V_{t+n^{2.5}}^{(i)}\mathbf{1}_{\mathcal{A}_{1,n^{2.5}}^{(i)}}|\mathcal{F}_t\big] \le ne^{-n}. \tag{6.55}$$

Next, by Theorem 5 of [9], there exists a constant $C = C(\varepsilon, c, d)$ that does not depend on n so that

$$\mathbb{E}\left[\sum_i \mathbf{1}_{\mathcal{A}_2^{(i)}}\right] \le C \tag{6.56}$$

for all n sufficiently large. By Corollary 5,

$$\mathbb{E}\left[\sum_i \mathbf{1}_{\mathcal{A}_3^{(i)}}\right] = O(\varepsilon)Y_t. \tag{6.57}$$

To bound $\mathcal{A}_4^{(i)}$, we consider the collection of colors J that are given in the definition of $\mathcal{A}_4^{(i)}$. We say that these colors are involved in an "unrecorded collision," as the collisions described in event $\mathcal{A}_4^{(i)}$ do not contribute to our count of the total number of near-collisions of the KCIP. Observe that any particle $R_s[j]$ with $j \in J$ that is involved in such an unrecorded collision at time s necessarily coalesces with another particle $R_s[j']$ during the course of the collision, and also must have $\tau_{\text{decoupling}}^{(j)} \le s$. Thus, each decoupled particle can be involved in only one unrecorded collision before being merged with another particle, and so

$$\sum_i \mathbf{1}_{\mathcal{A}_4^{(i)}} \le 2\left(\sum_i \mathbf{1}_{\mathcal{A}_{1,n^{2.5}}^{(i)}} + \sum_i \mathbf{1}_{\mathcal{A}_2^{(i)}} + \sum_i \mathbf{1}_{\mathcal{A}_3^{(i)}}\right). \tag{6.58}$$

Recall from equation (6.46) that $\mathcal{N}_s^{\text{near}}$ is the total number of near-collisions between times t and s. Combining inequalities (6.55), (6.56), (6.57) and (6.58), we have that

$$\mathbb{E}[\mathcal{N}_{t+\varepsilon n^3}^{\text{near}}] \geq Y_t(1 - O(\varepsilon)) - O(1). \tag{6.59}$$

Combining Proposition 6.20 with inequality (6.59), and noting that $n^{2.5} = o(n^3)$, completes the proof. □

Now, we are finally ready to give the proof of Theorem 6.1, establishing a drift condition for V_t.

PROOF OF THEOREM 6.1. Recall the definitions of $\delta > 0$ from Lemma 6.11 and $\kappa > 0$ from Lemma 6.15. Then

$$\begin{aligned}
&\mathbb{E}[V_{t+\varepsilon n^3}|\mathcal{F}_t] \\
&\quad = \mathbb{E}[V_{t+\varepsilon n^3} - Y_{t+\varepsilon n^3}|\mathcal{F}_t] + \mathbb{E}[Y_{t+\varepsilon n^3} - \tilde{Y}_{t+\varepsilon n^3}] + \mathbb{E}[\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t|\mathcal{F}_t] + \tilde{Y}_t \\
&\quad \leq 4cd(1 + o(1)) + \mathbb{E}[Y_{t+\varepsilon n^3} - \tilde{Y}_{t+\varepsilon n^3}|\mathcal{F}_t] + \mathbb{E}[\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t|\mathcal{F}_t] + \tilde{Y}_t \\
&\quad \leq 8cd(1 + o(1)) + \mathbb{E}[\tilde{Y}_{t+\varepsilon n^3} - \tilde{Y}_t|\mathcal{F}_t] + \tilde{Y}_t \\
&\quad \leq -\frac{2\delta}{3}\mathbb{E}[\mathcal{C}_{t+\varepsilon n^3}|\mathcal{F}_t] + O(\varepsilon\tilde{Y}_t) + O(1) + Y_t \\
&\quad \leq \left(1 - \frac{2\delta\kappa}{3}\right)Y_t + O(\varepsilon\tilde{Y}_t) + O(1) \\
&\quad \leq \left(1 - \frac{2\delta\kappa}{3} + O(\varepsilon)\right)V_t + O(1),
\end{aligned}$$

where the first inequality comes from Lemma 6.2, the second inequality comes from inequality (6.11) and a second application of Lemma 6.2, the third inequality comes from Lemma 6.11, the bound in the fourth inequality comes from Lemma 6.15, and the final inequality comes from the bound $Y_s \leq \tilde{Y}_s \leq V_s$ [see (6.11)].

Fixing $\varepsilon > 0$ sufficiently small, then, there exists some constant $C > 0$ so that

$$\mathbb{E}[V_{t+\varepsilon n^3}|\mathcal{F}_t] \leq \left(1 - \frac{\delta\kappa}{2}\right)V_t + C.$$

Write $\alpha = \frac{\delta\kappa}{2}$. Iterating, we have

$$\begin{aligned}
\mathbb{E}[V_{t+k\varepsilon n^3}] &\leq \mathbb{E}\big[\mathbb{E}\big[\cdots\mathbb{E}[V_{t+k\varepsilon n^3}|\mathcal{F}_{t+(k-1)\varepsilon n^3}]\cdots|\mathcal{F}_{t+\varepsilon n^3}\big]|\mathcal{F}_t\big] \\
&\leq \mathbb{E}\big[\mathbb{E}\big[\cdots\mathbb{E}\big[(1-\alpha)V_{t+(k-1)\varepsilon n^3} + C|t + (k-2)\varepsilon n^3\big]\cdots|t + \varepsilon n^3\big]|\mathcal{F}_t\big] \\
&\leq (1-\alpha)^k V_t + C_G
\end{aligned}$$

for some constant C_G and the proof is complete. □

7. Excursion lengths of KCIP. Fix ε_0 small enough so that Theorem 6.1 applies and set

$$k_{\max} = 4\frac{C_G}{\alpha}, \tag{7.1}$$

where C_G, α are as defined in inequality (6.1) of Theorem 6.1. In this section, the drift condition for V_t obtained in Theorem 6.1 will be used to show:

1. The distribution of the first hitting time of $\bigcup_{1\le k\le k_{\max}} \Omega_k$ is $O(n^3\log(n))$, uniformly in the starting point X_0 [see inequality (7.3)].
2. $T \gg n^3\log(n)$ implies that

$$\sum_{t\le T} \mathbf{1}_{X_t\in\bigcup_{1\le k\le k_{\max}}\Omega_k} \gg n^3\log(n)$$

with high probability, uniformly in the starting point X_0 (see Corollary 6).

Items (1) and (2) above thus provide strong bounds for the occupation times of KCIP on Ω_k, uniformly in $k \le k_{\max}$. We start with the following lemma (see Lemma 6.3 of [29] for a proof).

LEMMA 7.1. *Fix $0 < \beta < 1$ and $0 < \gamma < \infty$. Consider a stochastic process $\{J_t\}_{t\in\mathbb{N}}$ on $\mathbb{N}$ with associated filtration $\mathcal{J}_t$ that satisfies the drift condition*

$$\mathbb{E}[J_{s+1}|\mathcal{J}_s] \le (1-\beta)J_s + \gamma$$

for all $s \in \mathbb{N}$. Let $Z_1, Z_2, \ldots$ be an i.i.d. sequence of random variables with geometric distribution and mean $\frac{2}{\beta}$. If $J_0 \le \frac{4\gamma}{\beta}$, for all $T \in \mathbb{N}$ we have

$$\mathbb{P}\left[\sum_{s=0}^{T} \mathbf{1}_{J_s<\frac{4\gamma}{\beta}} < C\right] \le \mathbb{P}\left[\sum_{i=1}^{C} Z_i > T\right].$$

Define the set

$$\mathcal{K} = \left\{x \in \{0,1\}^{\Lambda(L,d)} : \sum_{v\in\Lambda(L,d)} x[v] \le k_{\max}\right\}. \tag{7.2}$$

We apply Lemma 7.1 to the KCIP on $\Lambda(L,d)$ $\lceil \varepsilon n^3\rceil$ times, with $J_s^{(i)} = V_{s\lceil\varepsilon n^3\rceil+i}$ and $\mathcal{J}_s = \mathcal{F}_{s\lceil\varepsilon n^3\rceil+i}$ for $0 \le i \le \lceil\varepsilon n^3\rceil$, to obtain the following corollary.

COROLLARY 6. *Fix ε sufficiently small so that Theorem 6.1 applies, and let α, C_G be as in Theorem 6.1. For fixed C_2 and $C_1 > \frac{16}{\alpha}C_2$ sufficiently large, all $n > N(c,d)$ sufficiently large, and any starting point $X_0 \in \Omega$,*

$$\mathbb{P}\left[\sum_{t=0}^{C_1 n^3\log(n)} \mathbf{1}_{X_t\in\mathcal{K}} < C_2 n^3\log(n)\right] = O\left(n^{-5}\right).$$

PROOF. Let $\tau_{\text{start}} = \inf\{t > 0 : X_t \in \mathcal{K}\}$ and fix $k \in \mathbb{N}$. By Theorem 6.1,

$$\mathbb{E}[V_{k\varepsilon n^3} \mathbf{1}_{\tau_{\text{start}} > k\varepsilon n^3}] \le \left(1 - \frac{1}{2}\alpha\right)^k V_0,$$

and so by Markov's inequality and the trivial bound that $V_t \le n$ for all t,

$$\begin{aligned} \mathbb{P}\big[\tau_{\text{start}} > k\varepsilon n^3\big] &\le \mathbb{P}[V_{k\varepsilon n^3} \mathbf{1}_{\tau_{\text{start}} > k\varepsilon n^3} > 1] \\ &\le n\left(1 - \frac{1}{2}\alpha\right)^k. \end{aligned} \tag{7.3}$$

Fix $T \in \mathbb{N}$ and let $\{Z_i\}_{i\in\mathbb{N}}$ be an i.i.d. sequence of random variables with geometric distribution and mean $\frac{2}{\alpha}$. By inequality (7.3), the Markov property and Lemma 7.1,

$$\begin{aligned} &\mathbb{P}\left[\sum_{t=0}^{C_1 n^3 \log(n)} \mathbf{1}_{X_t \in \mathcal{K}} > C_2 n^3 \log(n)\right] \\ &\quad \ge \mathbb{P}\left[\sum_{t=0}^{C_1 n^3 \log(n)} \mathbf{1}_{X_t \in \mathcal{K}} > C_2 n^3 \log(n) \Big| \tau_{\text{start}} < T\right] \mathbb{P}[\tau_{\text{start}} < T] \\ &\quad = \mathbb{P}[\tau_{\text{start}} < T] \sum_{t=0}^{T} \mathbb{P}\left[\sum_{t=0}^{C_1 n^3 \log(n)} \mathbf{1}_{X_t \in \mathcal{K}} > C_2 n^3 \log(n) \Big| \tau_{\text{start}} = t\right] \\ &\qquad \times \mathbb{P}[\tau_{\text{start}} = t | \tau_{\text{start}} \le T] \\ &\quad \ge \left(1 - n\left(1 - \frac{1}{2}\alpha\right)^{\lfloor \frac{T}{\varepsilon n^3} \rfloor}\right) \mathbb{P}\left[\sum_{t=T}^{C_1 n^3 \log(n)} \mathbf{1}_{X_t \in \mathcal{K}} > C_2 n^3 \log(n) \Big| \tau_{\text{start}} \le T\right] \\ &\quad \ge \left(1 - n\left(1 - \frac{1}{2}\alpha\right)^{\lfloor \frac{T}{\varepsilon n^3} \rfloor}\right) \left(1 - \lceil \varepsilon n^3 \rceil \mathbb{P}\left[\sum_{j=1}^{C_2 \log(n)} Z_j \le C_1 \log(n) - \frac{T}{\varepsilon n^3}\right]\right). \end{aligned}$$

Choosing $T = \lfloor \frac{C_1}{2} n^3 \log(n) \rfloor$, we have for C_1 sufficiently large that

$$\mathbb{P}\left[\sum_{j=1}^{C_2 \log(n)} Z_j \le C_1 \log(n) - \frac{T}{\varepsilon n^3}\right] = o(n^{-8}).$$

This completes the proof. □

7.1. *Bounds for the collision times of coalescent process.* Recall $k_{\max}$ from equation (7.1) and the set Ω_k from formula (2.1). In this section, we obtain estimates for the collision times for a coalescent process (see Definition 6.13) started in Ω_k. All of our bounds are based on soft arguments and are immediate consequences of results from [9]. These bounds will be used for obtaining estimates of occupation times of KCIP in Ω_k in the next section.

Let $\{Z_t\}_{t\in\mathbb{N}}$ be a coalescent process on $\Lambda(L,d)$ with $k_{\max}$ particles and moving rate $q=\frac{cd}{n^2}$. Define the process W_t on $\{0,1\}^{\Lambda(L,d)}$:

$$W_t[v]=\mathbf{1}_{\exists i:Z_t[i]=v}. \tag{7.4}$$

W_t will often be referred to as "the" coalescent process, as it is a Markov chain and $\{Z_t\}_{t\in\mathbb{N}}$ can be reconstructed (up to permutation of labels) from $\{W_t\}_{t\in\mathbb{N}}$. We give some notation related to the "skeletons" of our processes of interest. Define the sequence of times $\phi_0=0$ and

$$\phi_{i+1}=\inf\{t>\phi_i : W_t\neq W_{\phi_i}\}. \tag{7.5}$$

These are the times that $\{W_t\}_{t\in\mathbb{N}}$ changes.

REMARK 7.2. Let $\{\phi_i'\}_{i\geq 0}$ be a sequence of i.i.d. geometric random variables with mean 4 and define $\{W_t'\}_{t\geq 0}$ by

$$W_t'=W_{\phi_i}$$

for t satisfying $\sum_{j=0}^{i}\phi_j'\leq t<\sum_{j=0}^{i+1}\phi_j'$. The process $\{W_t'\}_{t\geq 0}$ is still a coalescent process in the sense of [9]. Furthermore,

$$\{W_{\phi_i}\}_{i\in\mathbb{N}}\overset{D}{=}\{W'_{\phi_i'}\}_{i\in\mathbb{N}} \tag{7.6}$$

and

$$\mathbb{E}[\phi_1]=\frac{n^2}{4k_{\max}cd}\mathbb{E}[\phi_1'].$$

Define the first collision time for the coalescent process as

$$\tau_{\mathrm{col}}=\inf\{t:|W_t|<|W_0|\}. \tag{7.7}$$

For the remainder of this section, define for $\zeta>0$:

$$\mathcal{G}_\zeta^{(n)}=\Big\{w\in\{0,1\}^{\Lambda(L,d)}:\inf_{u,v:w[u]=w[v]=1}|u-v|>\zeta\Big\}. \tag{7.8}$$

We show that, for ζ sufficiently large, collision times are not "too small" when started from $\mathcal{G}_\zeta^{(n)}\cap\Omega_k$.

LEMMA 7.3. *Let $\{Z_t\}_{t\in\mathbb{N}}$ be a coalescent process with $1<k\leq k_{\max}$ initial particles on graph $G=\Lambda(L,d)$ and let $\{W_t\}_{t\in\mathbb{N}}$ be defined as in formula (7.4). Then, for all $0<\delta<1$, there exist $\varepsilon=\varepsilon(c,d,k_{\max},\delta)>0$ and $C=C(c,d,k_{\max},\delta)$ so that, for all $n>N(c,d,k_{\max},\delta)$,*

$$\begin{aligned}\mathbb{P}[\tau_{\mathrm{col}}<\varepsilon n^3|W_0=w]&<1-\delta,\\ \mathbb{P}[\tau_{\mathrm{col}}<\phi_{\varepsilon n}|W_0=w]&<1-\delta\end{aligned} \tag{7.9}$$

uniformly in $1<k\leq k_{\max}$ and $w\in\mathcal{G}_C^{(n)}\cap\Omega_k$.

PROOF. We begin by proving the first half of inequality (7.9). By Theorem 5 of [9], for any fixed $\varepsilon > 0$, any sequence α_n satisfying $\lim_{n\to\infty} \alpha_n = \infty$, and any sequence $w^{(n)} \in \mathcal{G}_{\alpha_n}^{(n)}$, we have

$$\lim_{n\to\infty} \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n^3 | W_0 = w^{(n)}\big] = f(\varepsilon, c, d, k) < 1$$

for some explicit function f that satisfies $\lim_{\varepsilon\to 0} f(\varepsilon, c, d, k) = 0$ for $d \geq 3$ and all c, k. Defining $f(\varepsilon, c, d) = \max_{2\leq k\leq k_{\max}} f(\varepsilon, c, d, k)$, this implies

$$\lim_{n\to\infty} \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n^3 | W_0 = w^{(n)}\big] \leq f(\varepsilon, c, d) < 1 \tag{7.10}$$

and that $\lim_{\varepsilon\to 0} f(\varepsilon, c, d) = 0$ for all $d \geq 3$ and all c.

The remainder of the argument is a proof by contradiction. Fix $\delta > 0$ and choose $\varepsilon > 0$ so that $f(\varepsilon, c, d) \leq \frac{1-\delta}{2}$. Assume that inequality (7.9) is false. Then for all $C > 0$, there exists a strictly increasing sequence of integers $\{n_i = n_i(C)\}_{i\in\mathbb{N}}$ so that

$$\sup_{w^{(n_i)}\in\mathcal{G}_C^{(n_i)}} \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n_i^3 | W_0 = w^{(n_i)}\big] \geq 1 - \delta.$$

Let $\{C_j\}_{j\in\mathbb{N}}$ be a sequence satisfying $\lim_{j\to\infty} C_j = \infty$, and for each j let the increasing sequence of integers $\{n_{i,j} = n_i(C_j)\}_{i\in\mathbb{N}}$ satisfy

$$\sup_{w^{(n_{i,j})}\in\mathcal{G}_{C_j}^{(n_{i,j})}} \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n_{i,j}^3 | W_0 = w^{(n_{i,j})}\big] \geq 1 - \delta.$$

Then the diagonal sequence $n_{i,i}$ satisfies

$$\liminf_{i\to\infty} \sup_{w^{(n_{i,i})}\in\mathcal{G}_{C_i}^{(n_{i,i})}} \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n_{i,i}^3 | W_0 = w^{(n_{i,i})}\big] \geq 1 - \delta.$$

Since $f(\varepsilon, c, d) \leq \frac{1-\delta}{2}$ by assumption, this contradicts equality (7.10), completing the proof of the first half of inequality (7.9). The proof of the second half of inequality (7.9) is essentially identical; simply follow the same steps for the sped-up process $\{W'_t\}_{t\in\mathbb{N}}$ described in Remark 7.2 and use formula (7.6) to relate this back to $\{W_t\}_{t\in\mathbb{N}}$. □

Lemma 7.3 has the following strengthening as an immediate corollary.

COROLLARY 7. *There exist $\varepsilon = \varepsilon(c, d, k_{\max}) > 0$ and $\delta = \delta(c, d, k_{\max}) > 0$ so that, for all n sufficiently large,*

$$\begin{aligned} &\mathbb{P}\big[\tau_{\text{col}} < \varepsilon n^3 | W_0 = w\big] \leq 1 - \delta, \\ &\mathbb{P}[\tau_{\text{col}} < \phi_{\varepsilon n} | W_0 = w] \leq 1 - \delta \end{aligned} \tag{7.11}$$

uniformly in $1 < k \leq k_{\max}$ and $w \in \Omega_k$.

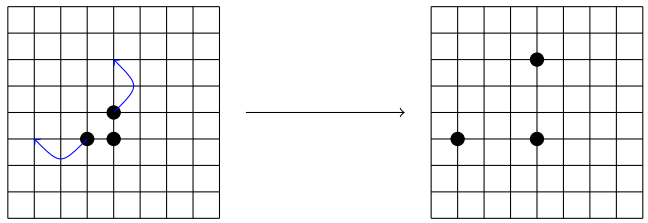

FIG. 7. *A path from a configuration with particles at distance* 1 *to a configuration with particles at distance* 3. *This path would require* 2 *particles to move* 2 *steps each.*

PROOF. Define the directed graph $\mathcal{X}$ to have vertex and directed edge sets

$$\begin{aligned} V(\mathcal{X}) &= \mathcal{K}, \\ E(\mathcal{X}) &= \big\{(x, y) : x, y \in \mathcal{K}, \mathbb{P}[W_1 = y | W_0 = x] > 0\big\}. \end{aligned} \tag{7.12}$$

Our key observation here is that it is possible to change any initial configuration $w \in \mathcal{K}$ into a configuration $w' \in \mathcal{G}_C^{(n)}$ by making some number of moves along the edges of $\mathcal{X}$, where the number of moves required is uniformly bounded in n (but may depend on C, d or $k_{\max}$). Since each move takes $O(n^2)$ steps in the coalescent process, and each move has probability bounded from below; this allows us to apply Lemma 7.3 to any initial configuration.

More formally, we begin by proving the first half of inequality (7.11). Fix $w \in \Omega_k$, any $0 < \gamma < 1$ and let $\varepsilon = \varepsilon(c, d, k_{\max}, \gamma)$, $C = C(c, d, k_{\max}, \gamma)$ be constants given by Lemma 7.3. Overloading notation slightly, if $x, y \in \mathcal{X}$, we denote by $|x - y|$ the length of the shortest path from x to y in $\mathcal{X}$; see Figure 7 for an example of such a path when $d = 2$ and $C = 2$.

Set

$$w' = \operatorname{argmin}\big\{w - u| : u \in \mathcal{G}_C^{(n)} \cap \Omega_k\big\}.$$

Let $\Gamma = (w_0 = w, w_1, \ldots, w_{|w-w'|} = w')$ be a path from w to w' satisfying $|w_i - w_{i+1}| = 1$ for all i. We have

$$\begin{aligned} \mathbb{P}\big[(W_{\phi_0}, W_{\phi_1}, \ldots, W_{\phi_{|w-w'|}}) = \Gamma | W_0 = w\big] &\geq (2dk_{\max})^{-|w-w'|} \\ &\geq (2dk_{\max})^{-k_{\max}^2(C+1)}. \end{aligned} \tag{7.13}$$

Thus, applying Lemma 7.3,

$$\begin{aligned} \mathbb{P}\big[\tau_{\text{col}} > \varepsilon n^3 | W_0 = w\big] &\geq \mathbb{P}\big[\tau_{\text{col}} > \varepsilon n^3 | W_0 = w' \in \mathcal{G}_C^{(n)}\big](2dk_{\max})^{-k_{\max}^2(C+1)} \\ &\geq \gamma (2dk_{\max})^{-k_{\max}^2(C+1)}. \end{aligned}$$

This completes the proof of the first half of inequality (7.11), with $\delta = \gamma(2dk_{\max})^{-k_{\max}^2(C+1)}$. As with Lemma 7.3, the proof of the second half is essentially identical. □

We further strengthen this to bound the probability that particles are close to colliding after a small number of steps. Let

$$\tau_{\text{near},i} = \inf\Big\{t : \inf_{u,v:W_t[u]=W_t[v]=1} |u-v| \le i\Big\}$$

be the first time that two particles in the coalescent process are within distance i.

LEMMA 7.4. *For all $i \in \mathbb{N}$ and all $0 < \gamma < 1$, there exist $\varepsilon = \varepsilon(c,d,k_{\max},\gamma,i) > 0$ and $C = C(c,d,k_{\max},\gamma,i) < \infty$ so that, for all $n > N(c,d,k_{\max},\gamma,i)$ sufficiently large,*

$$\begin{aligned} &\mathbb{P}\big[\tau_{\text{near},i} < \varepsilon n^3 | W_0 = w\big] \le 1-\gamma, \\ &\mathbb{P}[\tau_{\text{near},i} < \phi_{\varepsilon n} | W_0 = w] \le 1-\gamma \end{aligned} \tag{7.14}$$

uniformly in $1 < k \le k_{\max}$ and $w \in \mathcal{G}_C^{(n)}$.

PROOF. We begin by proving the first half of inequality (7.14). Fix $0 < \gamma < 1$, let $\delta = 1 - \frac{\gamma}{(4dk_{\max})^i}$, let C, ε be the constants associated with δ as given by Lemma 7.3, and let $w \in \mathcal{G}_C^{(n)}$. By the definition of $\tau_{\text{near},i}$, there is a sequence $\Gamma = (w_0 \equiv W_{\tau_{\text{near},i}}, w_1, \dots, w_i)$ so that $|w_{j+1} - w_j| = 1$ and so that a collision occurs during the transition from w_{i-1} to w_i. Let $I \in \mathbb{N}$ be such that $\phi_I = \tau_{\text{near},i}$. As in inequality (7.13),

$$\mathbb{P}\big[(W_{\phi_I}, W_{\phi_{I+1}}, \dots, W_{\phi_{I+i}}) = \Gamma | W_{\phi_I} = w_0\big] \ge (2dk_{\max})^{-i}.$$

Since

$$\mathbb{P}\big[\phi_{I+i} - \phi_I \ge n^2 \log(n)\big] = o(1),$$

this implies

$$\mathbb{P}\big[\tau_{\text{col}} < \tau_{\text{near},i} + n^2\log(n)\big] \ge \frac{1}{(2dk_{\max})^i}\big(1-o(1)\big)$$

uniformly in starting position W_0. Thus, for all $\varepsilon > 0$ and $w \in \mathcal{G}_C^{(n)}$,

$$\begin{aligned} \frac{\gamma}{(4dk_{\max})^i} &\ge \mathbb{P}\big[\tau_{\text{col}} < \varepsilon n^3 | W_0 = w\big] \\ &\ge \mathbb{P}\Big[\tau_{\text{col}} < \varepsilon n^3, \tau_{\text{near},i} < \frac{\varepsilon}{2} n^3 \Big| W_0 = w\Big] \\ &\ge \big(1-o(1)\big)(2dk_{\max})^{-i} \mathbb{P}\Big[\tau_{\text{near},i} < \frac{\varepsilon}{2} n^3 \Big| W_0 = w\Big]. \end{aligned}$$

We conclude that

$$\mathbb{P}\Big[\tau_{\text{near},i} < \frac{\varepsilon}{2} n^3 \Big| W_0 = w\Big] \le \frac{\gamma}{2}\big(1+o(1)\big),$$

completing the proof of the first half of inequality (7.14). As with Lemma 7.3, the proof of the second half is essentially identical. □

We then have the following Corollary to Lemma 7.4.

COROLLARY 8. *Fix $i \in \mathbb{N}$. There exist $\varepsilon = \varepsilon(c, d, k_{\max}, i) > 0$ and $\delta = \delta(c, d, k_{\max}, i) > 0$ so that, for all n sufficiently large,*

$$\mathbb{P}\big[\tau_{\text{near},i} < \varepsilon n^3 | W_0 = w\big] \leq 1 - \delta,$$
$$\mathbb{P}[\tau_{\text{near},i} < \phi_{\varepsilon n} | W_0 = w] \leq 1 - \delta$$

uniformly in $1 < k \leq k_{\max}$ and $w \in \Omega_k \cap \mathcal{G}_i^{(n)}$.

PROOF. This follows from Lemma 7.4 in essentially the same way that Corollary 7 followed from Lemma 7.3. □

7.2. *Comparing KCIP with the coalescent process.* Recall $k_{\max}$ from formula (7.1), the set Ω_k from formula (2.1) and the sets $\mathcal{G}_\zeta^{(n)}$ from formula (7.8). In this section, we obtain bounds on the occupation measure of X_t in Ω_k, uniformly in $k \leq k_{\max}$. Recall the definition of the exit times $L_k(x)$ from equation (2.2). Our intermediate steps are to show, uniformly in $1 \leq k \leq k_{\max}$:

1. Uniformly in $x \in \Omega_k$, $\mathbb{P}[L_k(x) < \varepsilon n^3] < 1 - \delta < 1$ for some $\varepsilon = \varepsilon(c, d, k_{\max})$ and $\delta = \delta(c, d, k_{\max})$.
2. Uniformly in $x \in \Omega_k$, $\mathbb{P}[L_k(x) > \varepsilon^{-1} n^3] < 1 - \delta$.
3. Uniformly in $x \in \Omega_k$, $\mathbb{P}[\sum_{t=0}^{\varepsilon n^3} \mathbf{1}_{X_t \in \Omega_k} < \frac{1}{2}\varepsilon n^3] < 1 - \delta$.
4. Uniformly in $x \in \Omega_k$, $\mathbb{P}[X_{L_k(X_0)} \in \Omega_{k-1} | X_0 = x] > \delta$, $\mathbb{P}[X_{L_k(X_0)} \in \Omega_{k+1} | X_0 = x] > \delta$.

We begin by proving item (4).

LEMMA 7.5. *There exist $\varepsilon = \varepsilon(c, d, k_{\max}) > 0$, $\delta = \delta(c, d, k_{\max}) > 0$ and $\alpha = \alpha(c, d, k_{\max}) > 0$ so that, uniformly in $1 \leq k \leq k_{\max}$ and $x \in \Omega_k$, we have*

$$\mathbb{P}\left[\sum_{t=0}^{\varepsilon n^3} \mathbf{1}_{X_t \in \Omega_k} \geq \alpha n^3 \,\middle|\, X_0 = x\right] \geq \delta.$$

PROOF. Assume first that $x \in \Omega_k \cap \mathcal{G}_2^{(n)}$. We couple $\{X_t\}_{t \in \mathbb{N}}$ to a coalescent process $\{Z_t\}_{t \in \mathbb{N}}$ using the same coupling as in Section 6.5.3 and define $\{W_t\}_{t \in \mathbb{N}}$ as in formula (7.4); since $X_0 \in \Omega_k$, we have $\tau_{\text{start}}^{(i)} = 0$ for all $1 \leq i \leq k$. Recalling the definition of $\{\phi_i\}_{i \in \mathbb{N}}$ from formula (7.5), we fix $T \in \mathbb{N}$ and consider the two events:

1. $\mathcal{A}_{1,T}$: The coalescent process has $\tau_{\text{near},2} < \phi_T$.

2. $\mathcal{A}_{2,T}$: The KCIP has $\inf\{t : \delta_t > 1\} < \phi_T$.

By Lemma 7.4, there exists some $\varepsilon_1, \gamma > 0$ so that

$$\mathbb{P}[\mathcal{A}_{1,\varepsilon n}] \le 1 - \gamma + o(1) \tag{7.15}$$

for $0 < \varepsilon < \varepsilon_1$. By Corollary 5, there exists some ε_2 so that for $0 < \varepsilon < \varepsilon_2$,

$$\mathbb{P}[\mathcal{A}_{2,\varepsilon n}] \le \frac{\gamma}{2} + o(1). \tag{7.16}$$

Choose $\varepsilon = \frac{1}{2}\min(\varepsilon_1, \varepsilon_2)$ and fix $\alpha, \beta > 0$. Set $\mathcal{A}_{\varepsilon n} = \mathcal{A}_{1,\varepsilon n} \cup \mathcal{A}_{2,\varepsilon n}$. By inequalities (7.15) and (7.16),

$$\begin{aligned}\mathbb{P}\left[\sum_{t=0}^{\beta n^3} \mathbf{1}_{X_t \in \Omega_k} \ge \alpha n^3 \Big| X_0 = x\right] &\ge \mathbb{P}\left[\sum_{t=0}^{\beta n^3} \mathbf{1}_{X_t \in \Omega_k} \ge \alpha n^3 \Big| X_0 = x, \mathcal{A}_{\varepsilon n}^c\right] \mathbb{P}[\mathcal{A}_{\varepsilon n}^c] \\ &\ge \frac{\gamma}{2}\mathbb{P}\left[\sum_{t=0}^{\beta n^3} \mathbf{1}_{X_t \in \Omega_k} \ge \alpha n^3 \Big| X_0 = x, \mathcal{A}_{\varepsilon n}^c\right] + o(1).\end{aligned}$$

Conditionally on W_{ϕ_i}, $\phi_{i+1} - \phi_i$ has geometric distribution with mean between $\frac{n^2}{cdk_{\max}}$ and $\frac{n^2}{cd}$. Thus, for all β sufficiently large,

$$\mathbb{P}\big[\beta^{-1} n^3 \le \phi_{\varepsilon n} \le \beta n^3\big] = 1 - o(1). \tag{7.17}$$

Since $\mathbb{P}[\mathcal{A}_{\varepsilon n}^c] \ge \frac{\gamma}{4} + o(1)$ is bounded away from 0, this implies that, for β sufficiently large,

$$\begin{aligned}&\mathbb{P}\left[\sum_{t=0}^{\beta n^3} \mathbf{1}_{X_t \in \Omega_k} \ge \alpha n^3 \Big| X_0 = x\right] \\ &\quad\ge \frac{\gamma}{4}\mathbb{P}\left[\sum_{t=0}^{\phi_{\varepsilon n}} \mathbf{1}_{X_t \in \Omega_k} \ge \alpha n^3 \Big| X_0 = x, \mathcal{A}_{\varepsilon n}^c\right] + o(1).\end{aligned} \tag{7.18}$$

By checking allowed sequences of update variables $\{v_s, p_s\}_{s \in \mathbb{N}}$, if $V_t \le k_{\max}$ and $\delta_t = 0$, we have

$$\begin{aligned}&\mathbb{P}\big[\delta_{t+1} = 1 | \delta_t = 0, V_t \le k_{\max}, t \le \phi_{\varepsilon n}, \mathcal{A}_{\varepsilon n}^c\big] \\ &\quad\le \mathbb{P}[\delta_{t+1} = 1 | \delta_t = 0, V_t \le k_{\max}] \\ &\quad\le \frac{2cdk_{\max}}{n^2},\end{aligned} \tag{7.19}$$

while if $V_t \le k_{\max} + 1$ and $\delta_t = 1$,

$$\begin{aligned}&\mathbb{P}\big[\delta_{t+1} = 0 | \delta_t = 1, V_t \le k_{\max} + 1, t \le \phi_{\varepsilon n}, \mathcal{A}_{\varepsilon n}^c\big] \\ &\quad\ge \mathbb{P}[\delta_{t+1} = 0 | \delta_t = 1, V_t \le k_{\max} + 1] \\ &\quad\ge \frac{1}{2n}\big(1 + o(1)\big).\end{aligned} \tag{7.20}$$

Since the lower bound in inequality (7.20) is $\Theta(n)$ times larger than the upper bound in inequality (7.19), we have

$$\mathbb{P}\left[\sum_{t=0}^{\phi_{\varepsilon n}} \mathbf{1}_{X_t \in \Omega_k} \geq \alpha \phi_{\varepsilon n} \Big| X_0 = x, \mathcal{A}_{\varepsilon n}^c\right] = 1 + o(1)$$

for all α^{-1} sufficiently large. Combining this with inequalities (7.17) and (7.18), we have for β, α^{-1} sufficiently large that

$$\mathbb{P}\left[\sum_{t=0}^{\beta n^3} \mathbf{1}_{X_t \in \Omega_k} \geq \alpha \beta^{-1} n^3 \Big| X_0 = x\right] \geq \frac{\gamma}{8} + o(1).$$

This completes the proof of the lemma when $x \in \Omega_k \cap \mathcal{G}_2^{(n)}$.

We give a short argument reducing the case $x \in \Omega_k$ to the case $x \in \Omega_k \cap \mathcal{G}_2^{(n)}$; it is essentially the same argument given in Corollary 7. Fix $x \in \Omega_k \setminus \mathcal{G}_2^{(n)}$. There exists $x' \in \Omega_k \cap \mathcal{G}_2^{(n)}$ and a sequence of configurations $\Gamma = (x = x_0, x_1, \ldots, x_\ell = x')$ so that $|x_{i+1} - x_i| \leq 1$ and $\ell \leq 16 k_{\max}^2$. The same argument as given in inequality (7.13) implies that there exists some $\gamma' = \gamma'(c, d, k_{\max}) > 0$ so that

$$\mathbb{P}\big[\inf\{t : X_t = x'\} < n^2 \log(n) | X_0 = x\big] > \gamma'.$$

This bound reduces the case $x \in \Omega_k$ to the case $x \in \Omega_k \cap \mathcal{G}_2^{(n)}$, at the cost only of replacing γ with $\gamma\gamma'$ and replacing εn^3 with $\varepsilon n^3 + n^2 \log(n) = (1 + o(1))\varepsilon n^3$. This completes the proof. □

Define

$$\rho_k = \inf\{t : X_t \in \Omega_k\}. \tag{7.21}$$

LEMMA 7.6. *For all $1 \leq k \leq k_{\max}$, there exists some $\delta, \varepsilon > 0$ so that*

$$\begin{aligned} &\mathbb{P}\big[\rho_{k-1} < \varepsilon n^3 | X_0 = x\big] > \delta, \\ &\mathbb{P}\big[\rho_{k+1} < \varepsilon n^3 | X_0 = x\big] > \delta \end{aligned} \tag{7.22}$$

holds uniformly in $x \in \Omega_k$ (where the first part of the inequality obviously requires $k \geq 2$).

PROOF. As argued in Corollary 7 (and again in the last paragraph of the proof of Lemma 7.5), we can assume without loss of generality that $x \in \Omega_k \cap \mathcal{G}_4^{(n)}$, since for any fixed $\varepsilon > 0$, the ratio

$$\frac{\inf_{x \in \Omega_k} \mathbb{P}[\rho_{k-1} < \varepsilon n^3 | X_0 = x]}{\inf_{x \in \Omega_k \cap \mathcal{G}_4^{(n)}} \mathbb{P}[\rho_{k-1} < \varepsilon n^3 (1 + o(1)) | X_0 = x]}$$

is bounded away from 0 uniformly in n. Thus, it is sufficient to consider only starting positions $X_0 \in \Omega_k \cap \mathcal{G}_4^{(n)}$.

Fix $1 \le k \le k_{\max}$ and $X_0 = x \in \Omega_k \cap \mathcal{G}_4^{(n)}$. We now couple $\{X_t\}_{t\in\mathbb{N}}$ to a coalescent process $\{Z_t\}_{t\in\mathbb{N}}$ using the same coupling as in Section 6.5.3 and define $\{W_t\}_{t\in\mathbb{N}}$ according to formula (7.4). Since $X_0 \in \Omega_k$, we have $\tau_{\text{start}}^{(i)} = 0$ for all $1 \le i \le k$. Define the sequence of movement times $\{\phi_j\}_{j\ge 0}$ as in formula (7.5) and the graph $\mathcal{X}$ as in formula (7.12). Fix $\varepsilon > 0$; we will define Ψ to be the collection of paths through the graph $\mathcal{X}$ for which the associated coalescent process has $\tau_{\text{near},4} \le \phi_{\varepsilon n}$. More precisely, we say that a path $\Gamma = (w_0, w_1, \dots, w_m)$ through the graph $\mathcal{X}$ is in Ψ if it satisfies:

1. $m \le \varepsilon n$.
2. $|w_{i+1} - w_i| = 1$ for all $0 \le i < m$, where distance is measured according to the graph distance on $\mathcal{X}$.
3. $w_0 = x$.
4. There exist $u, v \in G$ so that $|u - v| = 4$ and $w_m[u] = w_m[v] = 1$.
5. For all $u, v \in G$ with $|u - v| = 4$ and all $0 \le i < m$, we have $w_i[u]w_i[v] = 0$.

For any path $\Gamma \in \Psi$ of length m and any $A > 0$, we have by essentially the same calculation as in Lemma 4.1 that

$$\mathbb{P}\Big[\sup_{0\le t\le \tau_{\text{near},4}} \delta_t \le 1 | \{\phi_m > An^3\}, \{W_{\phi_i}\}_{i=0}^m = \Gamma\Big] \ge g(A) + o(1)$$

for some function $g(A) > 0$. Combining this with inequality (7.17), we have for all A so that $\frac{A}{\varepsilon}$ is sufficiently large that

$$\mathbb{P}\Big[\Big\{\sup_{0\le t\le \tau_{\text{near},4}} \delta_t \le 1\Big\} \cap \{\phi_m \le An^3\} | \{W_{\phi_i}\}_{i=0}^m = \Gamma\Big] \ge g(A) + o(1).$$

In particular, there exist $\varepsilon, A, B > 0$ so that

$$\mathbb{P}\Big[\Big\{\sup_{0\le t\le \tau_{\text{near},4}} \delta_t \le 1\Big\} \cap \{\phi_m \le An^3\} | \{W_{\phi_i}\}_{i=0}^m = \Gamma\Big] \ge B + o(1). \tag{7.23}$$

We choose such a triple ε, A, B for the remainder of the proof of the first half of inequality (7.22). Define $\rho'_{k-1} = \inf\{t : Y_t = k - 1\}$. By the same argument as in Proposition 6.20, there exists $\gamma' > 0$ so that

$$\mathbb{P}[\rho'_{k-1} < 4n^{2.5} | X_0 = y] \ge \gamma'(1 + o(1))$$

holds uniformly in $y \in \Omega_k \setminus \mathcal{G}_2^{(n)}$. By the same reduction argument used in Corollary 7, this implies that there exists some $\gamma > 0$ so that

$$\mathbb{P}[\rho'_{k-1} < 4n^{2.5} | X_0 = y] \ge \gamma(1 + o(1))$$

holds uniformly in $y \in \Omega_k \setminus \mathcal{G}_4^{(n)}$. Recall the definition of the update variables (p_t, v_t) from equation (1.1). Denote the event

$$\chi_n(t) = \Big\{\forall s : t \le s \le t + n\log(n)^2, p_s > \frac{c}{n} \text{ or } \sum_{(u,v_s)\in E(\Lambda(L,d))} X_s[u] = 0\Big\}. \tag{7.24}$$

We then have

$$\begin{aligned}
&\mathbb{P}\big[\rho_{k-1} < 4n^{2.5} + n\log(n)^2 | X_0 = y\big] \\
&\quad \geq \sum_{t=0}^{4n^{2.5}} \mathbb{P}\big[\rho_{k-1} < t + n\log(n)^2 | \rho'_{k-1} = t\big]\mathbb{P}\big[\rho'_{k-1} = t\big] \\
&\quad \geq \frac{1}{3}\sum_{t=0}^{4n^{2.5}} \mathbb{P}\left[\left\{\bigcup_{s=t}^{t+n\log(n)^2} \{v_s\} = \Lambda(L,d)\right\} \cap \chi_n(t) \,\middle|\, \rho'_{k-1} = t\right]\mathbb{P}\big[\rho'_{k-1} = t\big] \qquad (7.25) \\
&\quad \geq \frac{1}{3}\big(1 + o(1)\big)\sum_{t=0}^{4n^{2.5}} \mathbb{P}\big[\rho'_{k-1} = t\big] \\
&\quad \geq \frac{\gamma}{3}\big(1 + o(1)\big).
\end{aligned}$$

Combining this with inequality (7.23) and the observation that $w_m \in \Omega_k \setminus \mathcal{G}_4^{(n)}$ for all $\Gamma = (w_0, \ldots, w_m) \in \Psi$, we have

$$\mathbb{P}\big[\rho_{k-1} < An^3 + n^{2.5} + n\log(n)^2 | \{W_{\phi_i}\}_{i=0}^m = \Gamma\big] \geq \frac{B\gamma}{3} + o(1).$$

This implies that

$$\begin{aligned}
&\mathbb{P}\big[\rho_{k-1} < An^3 + n^{2.5} + n\log(n)^2 | X_0 = x\big] \\
&\quad \geq \sum_{\Gamma\in\Psi} \mathbb{P}\big[\rho_{k-1} < An^3 + n^{2.5} + n\log(n)^2 | \{W_{\phi_i}\}_{i=0}^m = \Gamma\big] \\
&\qquad \times \mathbb{P}\big[\{W_{\phi_i}\}_{i=0}^m = \Gamma | X_0 = x\big] \\
&\quad \geq \frac{B\gamma}{3}\sum_{\Gamma\in\Psi} \mathbb{P}\big[\{W_{\phi_i}\}_{i=0}^m = \Gamma | X_0 = x\big] + o(1) \\
&\quad = \frac{B\gamma}{3}\mathbb{P}[\tau_{\text{near},4} \leq \phi_{\varepsilon n} | X_0 = x] + o(1) \\
&\quad \geq \frac{B\gamma}{3}\mathbb{P}[\tau_{\text{col}} \leq \phi_{\varepsilon n} | W_0 = x] + o(1).
\end{aligned}$$

The last quantity, $\mathbb{P}[\tau_{\text{col}} \leq \phi_{\varepsilon n} | W_0 = x]$, is bounded away from 0 by Theorem 5 of [9]. This completes the proof of the first half of inequality (7.22).

We now prove the second half of inequality (7.22). Fix $x \in \mathcal{G}_4^{(n)}$ and $\varepsilon > 0$. We consider a new collection of paths Ψ on $\mathcal{X}$. Roughly speaking, these will be the paths for which there are no near-collisions for many steps. More precisely, we say that $\Gamma = (x = w_0, w_1, \ldots, w_m)$ is in Ψ if:

- $m \geq \frac{1}{2}\varepsilon n$.
- $|w_{i+1} - w_i| = 1$ for all $0 \leq i < m$.

- $w_0 = x$.
- For all $u, v \in G$ with $|u - v| \le 4$ and all $0 \le i \le m$, we have $w_i[u]w_i[v] = 0$.

Recall that $\tau^{(j)}_{\text{decoupling}}$ is defined in formula (6.44) and $\zeta^{(j)}_{\text{triple}}$ is defined in formula (6.41). If $\{W_{\phi_i}\}_{i=0}^m = \Gamma \in \Psi$, then no near-collisions of W_t have occurred by time ϕ_m, and so conditioned on this event we have

$$\min_{1\le j\le k} \tau^{(j)}_{\text{decoupling}} \ge \min\Big(\phi_m, \min_{1\le j\le k} \zeta^{(j)}_{\text{triple}}\Big). \tag{7.26}$$

Also, for any $t \in \mathbb{N}$ and any starting point $X_0 = x \in \Omega_k$ with $\delta_0 = 1$, we have

$$\mathbb{P}\Big[\min_{0\le s\le t} \delta_s \ge 1 \big| X_0 = x\Big] \ge \left(1 - \frac{2}{n}\right)^t. \tag{7.27}$$

In particular, the indicator function of the event $\{\min_{0\le s\le t} \delta_s \ge 1\}$ is stochastically dominated by a geometric random variable with mean $\frac{n}{2}$. Combining inequalities (7.26) and (7.27), we have

$$\begin{aligned}\mathbb{P}\left[\left\{\sum_{t=0}^{\phi_m} \mathbf{1}_{\delta_t \ge 1} \ge \frac{\varepsilon}{32} n^2\right\} \cup \left\{\min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} \le \phi_m\right\} \Big| \{W_{\phi_i}\}_{i=0}^m = \Gamma\right] \\ = 1 - o(1).\end{aligned} \tag{7.28}$$

Noting that

$$\mathbb{P}\Big[\min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} = t + 1 | \{W_{\phi_i}\}_{i=0}^m = \Gamma, \delta_t = 1, \min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} \ge t, t \le \phi_i\Big] \ge \frac{c}{n^2},$$

we obtain

$$\begin{aligned}&\mathbb{P}\left[\min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} > \phi_m | \{W_{\phi_i}\}_{i=0}^m = \Gamma, \sum_{t=0}^{\phi_m} \mathbf{1}_{\delta_t \ge 1} \ge \frac{\varepsilon}{32} n^2\right] \\ &\quad \le \left(1 - \frac{c}{n^2}\right)^{\frac{\varepsilon}{32} n^2} \\ &\quad \le e^{-\frac{\varepsilon c}{32}}.\end{aligned}$$

Combining this with inequality (7.28) and rearranging terms, this implies

$$\begin{aligned}&\mathbb{P}\Big[\min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} \le \phi_m | \{W_{\phi_i}\}_{i=0}^m = \Gamma\Big] \\ &\quad \ge 1 - o(1) - \mathbb{P}\left[\min_{1\le j\le k} \zeta^{(j)}_{\text{triple}} > \phi_m \big| \{W_{\phi_i}\}_{i=0}^m = \Gamma, \sum_{t=0}^{\phi_m} \mathbf{1}_{\delta_t \ge 1} \ge \frac{\varepsilon}{32} n^2\right] \\ &\quad \ge 1 - e^{-\frac{\varepsilon c}{32}} - o(1).\end{aligned}$$

Summing over $\Gamma \in \mathcal{X}$,

$$\begin{aligned}
&\mathbb{P}\Big[\min_{1\le j\le k} \zeta_{\text{triple}}^{(j)} \le \min(\phi_m, \tau_{\text{near},4})\Big] \\
&\quad \ge \sum_{\Gamma\in\mathcal{X}} \mathbb{P}\Big[\min_{1\le j\le k} \zeta_{\text{triple}}^{(j)} \le \phi_m \big| \{W_{\phi_i}\}_{i=0}^m = \Gamma\Big]\mathbb{P}\big[\{W_{\phi_i}\}_{i=0}^m = \Gamma\big] \\
&\quad \ge \big(1 - e^{-\frac{\varepsilon c}{32}} - o(1)\big) \sum_{\Gamma\in\mathcal{X}} \mathbb{P}\big[\{W_{\phi_i}\}_{i=0}^m = \Gamma\big] \\
&\quad \ge \big(1 - e^{-\frac{\varepsilon c}{32}} - o(1)\big)\mathbb{P}[\tau_{\text{near},4} \ge \phi_{\frac{\varepsilon}{2}}].
\end{aligned}$$

Write $\zeta_{\text{triple}}^{(\min)} = \min_{1\le j\le k} \zeta_{\text{triple}}^{(j)}$. By Corollary 8, we conclude that there exists some $\gamma > 0$ so that

$$\mathbb{P}\big[\zeta_{\text{triple}}^{(\min)} \le \min(\phi_m, \tau_{\text{near},4})\big] \ge \gamma + o(1).$$

Combining this with the calculation in inequality (7.25),

$$\begin{aligned}
&\mathbb{P}\big[\rho_{k+1} < \phi_{\frac{\varepsilon n}{2}} + n\log(n)^2 \big| X_0 = y\big] \\
&\quad \ge \mathbb{P}\Bigg[\Bigg\{\bigcup_{s=\zeta_{\text{triple}}^{(\min)}}^{\zeta_{\text{triple}}^{(\min)} + n\log(n)^2} \{v_s\} = \Lambda(L,d)\Bigg\} \\
&\qquad \cap \chi_n\big(\zeta_{\text{triple}}^{(\min)}\big) \big| \zeta_{\text{triple}}^{(\min)} \le \min(\phi_{\frac{\varepsilon}{2}n}, \tau_{\text{near},4}), X_0 = y\Bigg] \\
&\qquad \times \mathbb{P}\big[\zeta_{\text{triple}}^{(\min)} \le \min(\phi_{\frac{\varepsilon}{2}n}, \tau_{\text{near},4}) \big| X_0 = y\big] \\
&\quad \ge \big(1 - o(1)\big)\big(\gamma - o(1)\big).
\end{aligned}$$

Combining this with inequality (7.17) completes the proof. □

8. Proof of Theorem 2. In this section, we find bounds on the occupation times of Ω_k and use these bounds to finish the proof of Theorem 2. Recall from (3.2) that the quantity κ_k applied to X_t is

$$\kappa_k(T) = \sum_{t=0}^{T} \mathbf{1}_{X_t\in\Omega_k}.$$

LEMMA 8.1 (Bound on occupation measures). *There exists* $0 < \gamma < \infty$ *so that, for all* A *sufficiently large and all* $B > 0$,

$$\mathbb{P}\big[\kappa_k\big(An^3\log(n)\big) < Bn^3\log(n)\big] \le \gamma\frac{B}{A} + o(1)$$

uniformly in $X_0 = x \in \Omega$ *and* $1 \le k \le k_{\max}$.

PROOF. Recall the set $\mathcal{K}$ from (7.2) and $k_{\max}$ from (7.1). By Corollary 6, there exists $c_1 > 0$ so that for all A sufficiently large,

$$\mathbb{P}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\mathcal{K}} > c_1 An^3\log(n)\Big| X_0 = x \in \Omega\right] = 1 + o(1). \tag{8.1}$$

Recall the definition of the update variables (p_t, v_t) from formula (1.1) and $\chi_n \equiv \chi_n(0)$ from formula (7.24). For any $x \in \mathcal{K}$, we have that

$$\begin{aligned}
&\mathbb{P}\left[\inf\left\{t > 0 : X_t \in \mathcal{K} \cap \bigcup_k \Omega_k\right\} < n\log(n)^2 \Big| X_0 = x \in \mathcal{K}\right] \\
&\geq \mathbb{P}\left[\left\{\bigcup_{s=0}^{n\log(n)^2} \{v_s\} = \Lambda(L,d)\right\} \cap \chi_n \Big| X_0 = x \in \mathcal{K}\right] \qquad (8.2)\\
&= 1 + o(1).
\end{aligned}$$

Combining inequalities (8.1) and (8.2) with Lemma 7.5, there exists some $c_2 > 0$ and $1 \leq k' \leq k_{\max}$ so that, for all A sufficiently large,

$$\mathbb{P}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'}} > c_2 An^3\log(n)\Big| X_0 = x \in \Omega\right] = 1 + o(1). \tag{8.3}$$

Inequality (8.3) implies the existence of some $c_3 > 0$ so that, for any $X_0 = x \in \Omega$,

$$\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'}}\right] \geq c_3 An^3\log(n) \tag{8.4}$$

for all A sufficiently large. We claim that inequality (8.4) also holds with k' replaced by $k'+1$ and also (if $k' \geq 2$) with $k'-1$. By Lemmas 7.6 and 7.5, there exists some $c_4 > 0$ so that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'+1}}\right] \geq c_4\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'}}\right] \geq c_4c_3An^3\log(n), \\
&\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'-1}}\right] \geq c_4\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_{k'}}\right] \geq c_4c_3An^3\log(n).
\end{aligned} \tag{8.5}$$

This is exactly inequality (8.4) with k' replaced by $k'+1$ and $k'-1$, respectively. Since the set $\{1, 2, \ldots, k_{\max}\}$ is finite, the argument between inequalities (8.4) and (8.5) implies that there exists some $c_5 > 0$ so that

$$\mathbb{E}\left[\sum_{t=0}^{An^3\log(n)} \mathbf{1}_{X_t\in\Omega_j}\right] \geq c_5 An^3\log(n) \tag{8.6}$$

for all $1 \le j \le k_{\max}$. Since $\kappa_k(T) = \sum_{t=0}^{T} \mathbf{1}_{X_t \in \Omega_k}$, the result now follows from an application of Markov's inequality to inequality (8.6). $\square$

Finally, we prove our main result.

PROOF OF THEOREM 2. The lower bound in Theorem 2 is given in Theorem 3. We now show the proof of the upper bound in Theorem 2 by applying the bounds in Lemma 5.1 and Lemma 8.1 to Lemma 3.1.

Lemma 8.1 implies that there exist constants $0 < A, B < \infty$, and a function $N = N(A, B, c, d)$ so that for all $k \le k_{\max}$ [see formula (7.1)] and all $n > N$,

$$\sup_{x \in \Omega} \mathbb{P}\big[\kappa_k\big(An^3 \log(n)\big) < Bn^3 \log(n) | X_0 = x\big] \le \frac{1}{16}. \tag{8.7}$$

Similarly, Lemma 5.1 implies that there exists a constant $0 < C_1 < \infty$ so that

$$\max_{1 \le k \le k_{\max}} \tau_{n,k} \le C_1 n^{2+\frac{2}{d}} \log(n). \tag{8.8}$$

In the notation of Lemma 3.1, let K be the kernel of the KCIP, let $\Theta = \Omega$, let $\Theta_k = \Omega_k$ for $1 \le k \le \frac{n}{2}$ and let $\Theta_{\frac{n}{2}+1} = \Omega \setminus \bigcup_{k=1}^{\frac{n}{2}} \Omega_k$. Also set $I = \{1, 2, \ldots, k_{\max}\}$ and $S = \bigcup_{i=1}^{\lfloor \log(n) \rfloor} \Omega_k$. We note that

$$\begin{aligned} \pi(S) &= 1 - o(1), \\ \min_{x \in S} K[x, x] &= 1 - o(1), \\ \max_{x \notin S} \sum_{y \in S} K[x, y] &= 1 - o(1), \end{aligned}$$

and so our chain is *sufficiently lazy* for all $n > N_0(c, d)$ large.

By Theorem 6.1, if the KCIP is started at the stationary distribution π, there exists $0 < \alpha, \varepsilon < 1$ so that

$$\mathbb{E}[V_0] = \mathbb{E}[V_{\varepsilon n^3}] \le (1 - \alpha)\mathbb{E}(V_0) + C_G.$$

Thus, we get $\mathbb{E}(V_0) \le \frac{C_G}{\alpha} = \frac{k_{\max}}{4}$. Markov's inequality yields that

$$\mathbb{P}[V_0 \le k_{\max}] \ge \frac{3}{4}.$$

Since V_0 is distributed according to π, this immediately yields

$$\pi\left(\bigcup_{k \in I} \Omega_k\right) \ge \frac{3}{4} - o(1).$$

Thus, in the notation of Lemma 3.1, we can fix for all $n > N = N(c, d, A, B)$ sufficiently large,

$$a = \frac{3}{8}, \qquad \beta = \frac{11}{16}, \qquad \gamma = \frac{1}{11},$$
$$t = 16c'_\gamma C_1 n^{2+\frac{2}{d}} \log(n) \le Bn^3 \log(n),$$
$$T = \lceil An^3 \log(n) \rceil,$$

where C_1 is the constant from (8.8). By (8.8), we have

$$\max_{1 \le k \le k_{\max}} \frac{\tau_{n,k} c'_\gamma}{t} \le \frac{1}{16}. \tag{8.9}$$

By (8.7), we have that for $1 \le k \le k_{\max}$,

$$\sup_{x \in \Omega} \mathbb{P}\big[\kappa_k(T) < t | X_0 = x\big] \le \frac{1}{16}. \tag{8.10}$$

Combining inequalities (8.9) and (8.10) immediately implies that, in the notation of Lemma 3.1, $\mathcal{T} \le T = O(n^3 \log(n))$. Thus, by Lemma 3.1 we have that $\tau_{\text{mix}} = O(n^3 \log(n))$ and the proof of Theorem 2 is complete. □

9. Conclusion and future work. We have resolved only some special cases in Aldous' conjecture. In particular, we do not have any results for:

- Graphs other than the torus.
- Density regimes other than $p = \frac{c}{n}$.

In this section, we give conjectures for other graphs and regimes, and mention cases for which our methods work well. Before doing so, there is the question as to whether the adjustment to Aldous' conjecture required by Theorem 3 is essentially the only required correction. If so, this would suggest the conjecture.

CONJECTURE 9. *Fix $d \in \mathbb{N}$ and let G be a d-regular graph. The mixing time τ_{mix} of the KCIP with parameter p on graph G is $O(p^{-1} d|G| \tau_{\text{mix}}^{\text{RW}} + \frac{|G|}{d^3 p^2})$ as $|G|$ goes to infinity, where $\tau_{\text{mix}}^{\text{RW}}$ is the mixing time of the $\frac{1}{2}$-lazy simple random walk on the graph G.*

We believe that this gives the correct answer for the torus in dimension $d \ge 3$ [i.e., we conjecture that the correct mixing time for the process described in Theorem 2 is $O(n^3)$]. However, we do not believe that Conjecture 9 is true in general. In particular, define the coalescence time τ_{Coal} of a graph to be the expected time for the coalescent process on G to go from $|G|$ particles to 1 particle. We suspect that the mixing time of the KCIP is bounded from below by the coalescence time of the associated coalescent process on the same graph. By [9], this suggests that

the mixing time for the torus in dimension $d = 2$ is at least $n^3 \log(n)$, while Conjectures 1 and 9 both suggest the mixing time is $O(n^3)$. In general, we conjecture the following.

CONJECTURE 10. *Fix $d \in \mathbb{N}$ and let G be a d-regular graph. The mixing time τ_{mix} of the KCIP with parameter p on graph G is $O(p^{-1}d|G|\tau_{\text{mix}}^{\text{RW}} + \frac{|G|}{d^3p^2} + p^{-1}d|G|\tau_{\text{Coal}})$ as $|G|$ goes to infinity, where $\tau_{\text{mix}}^{\text{RW}}$ is the mixing time of the $\frac{1}{2}$-lazy simple random walk on the graph G and τ_{Coal} is the coalescence time of the graph G.*

Next, we discuss when our strategy outlined in Section 2.2 may be applicable for proving Conjecture 10. When restricted to densities in the regime $p = \frac{c}{n}$, our proof strategies are likely to work well for many other sequences of bounded-degree graphs. In particular, for random triangle-free d-regular graphs, our argument goes through with only two major changes. The first change is to replace all bounds on the coalescent process from [9] with analogous bounds from [26] and [27]. These bounds are substantially looser, but strong enough for our arguments to go through. Next, random d-regular graphs are expanders with high probability (see, e.g., [18]), and in particular have spectral gaps that are uniformly bounded below in n and mixing times that grow like $\log(n)$. This allows us to make the second change, replacing the bound on the log-Sobolev constant from [32] with a bound on the spectral gap from [18]. Besides the invocation of [32], all comparison arguments for the log-Sobolev constant of $\Lambda(L, d)$ given in Section 5 also apply as written to bounding the spectral gap of general d-regular graphs. These graphs have such large spectral gaps that the bounds obtained this way are sufficient. Beyond expanders, we generally expect our strategy to succeed for families of bounded-degree graphs with

$$\max\bigl(\tau_{\text{Coal}}, \tau_{\text{mix}}^{\text{RW}}\bigr) = O\bigl(n^3\bigr). \tag{9.1}$$

In particular, a similar approach works for the lattice in $d \geq 3$ dimensions. When (9.1) fails to hold, as with $\Lambda(L, d)$ in dimension $d = 2$, the arguments in Section 6 must be substantially changed.

For sequences of $m = m(n)$-regular graphs [with $m(n)$ very slowly growing with n], our strategy could be refined to give nontrivial bounds, but our results are not very useful as written. For graphs with very large degrees $d \approx n$, a straightforward comparison argument to the KCIP on the complete graph analogous to that given in [10] for the simple exclusion process gives useful bounds.

Our strategies are unlikely to work well for $p = p_n \gg \frac{c}{n}$, as the arguments this paper rely quite strongly on the stationary measure of the constrained Ising process being concentrated on configurations with few particle. Despite technical difficulties, we believe that the $O(n^3)$ bound will hold up to $p_n \approx n^{-\frac{1}{2}-\frac{1}{d}}$. Again, for p sufficiently large, a straightforward comparison argument to the constrained Ising process on the complete graph analogous to that given in [10] for the simple exclusion process gives useful bounds.

Acknowledgements. We thank Daniel Jerison and Anastasia Raymer for helpful conversations, and a referee for a very careful reading.

REFERENCES

[1] ALDOUS, D. Open Problem List. http://www.stat.berkeley.edu/~aldous/Research/OP/. Accessed: 2014-12-18.

[2] ALDOUS, D. and DIACONIS, P. (2002). The asymmetric one-dimensional constrained Ising model: Rigorous results. *J. Stat. Phys.* **107** 945–975. MR1901508

[3] CANCRINI, N., MARTINELLI, F., ROBERTO, C. and TONINELLI, C. (2009). Kinetically constrained models. *New Trends in Mathematical Physics* 741–752.

[4] CHLEBOUN, P., FAGGIONATO, A. and MARTINELLI, F. (2014). Time scale separation and dynamic heterogeneity in the low temperature East model. *Comm. Math. Phys.* **328** 955–993. MR3201217

[5] CHLEBOUN, P., FAGGIONATO, A. and MARTINELLI, F. (2015). Mixing time and local exponential ergodicity of the East-like process in $\mathbb{Z}^d$. *Ann. Fac. Sci. Toulouse Math.* (6) **24** 717–743. MR3434253

[6] CHLEBOUN, P., FAGGIONATO, A. and MARTINELLI, F. (2016). Relaxation to equilibrium of generalized East processes on $\mathbb{Z}^d$: Renormalization group analysis and energy-entropy competition. *Ann. Probab.* **44** 1817–1863.

[7] CHLEBOUN, P. and MARTINELLI, F. (2013). Mixing time bounds for oriented kinetically constrained spin models. *Electron. Commun. Probab.* **18** no. 60, 9. MR3084571

[8] CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika* **60** 581–588. MR0343950

[9] COX, J. T. (1989). Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17** 1333–1366. MR1048930

[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. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** 695–750. MR1410112

[12] DYER, M., GOLDBERG, L. A., JERRUM, M. and MARTIN, R. (2006). Markov chain comparison. *Probab. Surv.* **3** 89–111. MR2216963

[13] FREDRICKSON, G. and ANDERSEN, H. (1984). Kinetic Ising model of the glass transition. *Phys. Rev. Lett.* **53** 1244–1247.

[14] FREDRICKSON, G. and ANDERSEN, H. (1985). Facilitated kinetic Ising models and the glass transition. *J. Chem. Phys.* **83** 5822–5831.

[15] GANGULY, S., LUBETZKY, E. and MARTINELLI, F. (2015). Cutoff for the East process. *Comm. Math. Phys.* **335** 1287–1322. MR3320314

[16] GARRAHAN, J., SOLLICH, P. and TONINELLI, C. (2011). *Dynamical Heterogeneities in Glasses, Colloids and Granular Media*. Oxford Univ. Press, London.

[17] HOLLEY, R. A. and LIGGETT, T. M. (1975). Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** 643–663. MR0402985

[18] HOORY, S., LINIAL, N. and WIGDERSON, A. (2006). Expander graphs and their applications. *Bull. Amer. Math. Soc.* (*N.S.*) **43** 439–561 (electronic). MR2247919

[19] JACK, R. L., MAYER, P. and SOLLICH, P. (2006). Mappings between reaction–diffusion and kinetically constrained systems: $A + A \leftrightarrow A$ and the Fredrickson–Andersen model have upper critical dimension $d_c = 2$. *J. Stat. Mech. Theory Exp.* **2006** P03006.

[20] JERRUM, M., SON, J.-B., TETALI, P. and VIGODA, E. (2004). Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Probab.* **14** 1741–1765. MR2099650

[21] KORDZAKHIA, G. and LALLEY, S. P. (2006). Ergodicity and mixing properties of the northeast model. *J. Appl. Probab.* **43** 782–792. MR2274800

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

[23] MARTINELLI, F. (2014). East model: Mixing time, cutoff and dynamical heterogeneities. Lecture (Slides at http://www2.warwick.ac.uk/fac/sci/maths/research/events/2013-2014/nonsymp/gscsd/programme/fabio-slides.pdf).

[24] MARTINELLI, F. and TONINELLI, C. (2013). Kinetically constrained spin models on trees. *Ann. Appl. Probab.* **23** 1967–1987. MR3134727

[25] OLIVEIRA, R. I. (2012). Mixing and hitting times for finite Markov chains. *Electron. J. Probab.* **17** no. 70, 12. MR2968677

[26] OLIVEIRA, R. I. (2012). On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.* **364** 2109–2128. MR2869200

[27] OLIVEIRA, R. I. (2013). Mean field conditions for coalescing random walks. *Ann. Probab.* **41** 3420–3461. MR3127887

[28] PERES, Y. and SOUSI, P. (2013). Mixing times are hitting times of large sets. *J. Theoret. Probab.* **9** 459–510.

[29] PILLAI, N. and SMITH, A. (2015). Elementary bounds on mixing times for decomposable Markov chains. Preprint. Available at arXiv:1507.04028.

[30] SMITH, A. (2015). Comparison theory for Markov chains on different state spaces and application to random walk on derangements. *J. Theoret. Probab.* **28** 1406–1430.

[31] VALIANT, P. (2004). Linear bounds on the North–East model and higher-dimensional analogs. *Adv. in Appl. Math.* **33** 40–50. MR2064356

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

DEPARTMENT OF STATISTICS
HARVARD UNIVERSITY
1, OXFORD ST.
CAMBRIDGE, MASSACHUSETTS 01880
USA
E-MAIL: pillai@fas.harvard.edu

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF OTTAWA
585 KING EDWARD AVENUE
OTTAWA
ON K1N 7N5
CANADA
E-MAIL: asmi28@uottawa.ca