Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2019, Vol. 55, No. 3, 1649–1678
https://doi.org/10.1214/18-AIHP930

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Mixing times for a constrained Ising process on the two-dimensional torus at low density

Natesh S. Pillai[a] and Aaron Smith[b]

[a]*Department of Statistics, Harvard University, 1 Oxford Street, Cambridge MA 02138, USA. E-mail: pillai@fas.harvard.edu*
[b]*Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa ON K1N 7N5, Canada. E-mail: smith.aaron.matthew@gmail.com*

Received 29 August 2017; revised 24 May 2018; accepted 4 September 2018

Abstract. 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 locations of the particles. 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. In this note, we study the mixing time of a KCIP on the 2-dimensional torus $G = \mathbb{Z}_L^2$ in the low-density regime $p = \frac{c}{L^2}$ for arbitrary $0 < c < \infty$, extending our previous results for the analogous process on the torus $\mathbb{Z}_L^d$ in dimension $d \geq 3$. Our general approach is similar, but the extension requires more delicate bounds on the behaviour of the process at intermediate densities.

Résumé. Nous étudions un processus d'Ising avec contraintes cinétiques (PICC) associé à un graphe G et un paramètre de densité p. Ce processus est un système de particules en interaction avec espace d'états $\Omega = \{0,1\}^G$, décrivant les positions des particules. Les « contraintes » apparaissant dans le nom de ce processus réfèrent à la règle suivante: un sommet ne peut pas changer son état à moins qu'il ait un voisin dans l'état « 1 ». Le PICC a été proposé par des physiciens comme un modèle pour la transition vitreuse. Dans ce travail, nous analysons le temps de mélange d'un PICC sur le tore de dimension 2 $G = \mathbb{Z}_L^2$ dans le régime de faible densité $p = \frac{c}{L^2}$, où $0 < c < \infty$. Ceci prolonge nos résultats au processus analogue sur le tore $G = \mathbb{Z}_L^d$, $d \geq 3$. Notre approche générale est similaire, mais cette extension requiert des bornes plus délicates sur le comportement du processus aux densités intermédiaires.

MSC: Primary 60J10; secondary 60J20

Keywords: Kinetically constrained process; Interacting particle systems; Markov chain; Mixing time; Glass transition

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. These processes have also appeared outside of the computer science literature (see the surveys [7,8] for examples). In this paper, we analyze one of the simplest and most-studied processes introduced in [13,14], called the FA1f process. The FA1f process takes as parameters the underlying graph G and the typical density $0 < p < 1$ of 1's at equilibrium. The mixing time τ_{mix} of this process at small density $p = \frac{c}{|G|}$ for fixed $0 < c < \infty$ is the subject of a well-known conjecture of Aldous [1]:

$$\tau_{\text{mix}} \approx p^{-2} \tau_{\text{mix}}^{(G)},$$

where $\tau_{\text{mix}}^{(G)}$ is the mixing time of simple random walk on G. The conjecture is based on the heuristic that, near equilibrium, the FA1f process at low temperature behaves much like a simple random walk on G with roughly $p|G|$ walkers, slowed down by a factor of p^{-3}.

In previous work [32], we studied Aldous' conjecture in the case that the underlying graph is the torus $\mathbb{Z}_L^d$ in dimension $d \geq 3$. In that paper, we showed that Aldous' conjecture does not quite hold for these examples: while the heuristic is correct near equilibrium, the mixing time is governed by the much larger time it takes for the initial all-1's configuration to drift towards a more typical configuration with roughly c 1's. As we show in [32] in the special case of the torus in dimension $d \geq 3$, this drift time can be related to the time it takes coalescing random walks on the same underlying graph to coalesce. In this paper, we extend our previous work to the more difficult case of $d = 2$.

We recall the definition of the FA1f process on a general connected finite graph $G = (V, E)$ with 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 by the following update procedure. At each time $t \in \mathbb{N}$, 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 $p_t \leq p$ 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 $w \in V \setminus \{v_t\}$.

Set $|V| = n$; for general points $x \in \{0, 1, 2, \ldots\}^G$, define $|x| = \sum_{v\in G} \mathbf{1}_{x[v]\neq 0}$. Let π denote the stationary distribution of $\{X_t\}_{t\in\mathbb{N}}$. For $y \in \Omega$, this stationary distribution 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 normalizing constant. Thus $\pi(y)$ is proportional to the Binomial(n, p) distribution on the number of non-zero labels in $y \in \Omega$, conditional on having at least one non-zero entry.

We give some standard notation. 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

$$\tau(\varepsilon) = \inf\Big\{t > 0 : \sup_{X_1=x\in\Omega} \big\|\mathcal{L}(X_t) - \pi\big\|_{\mathrm{TV}} < \varepsilon\Big\}$$

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

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 $d = 2$.

The following is our main result:

Theorem 1 (Mixing of the constrained Ising process on the torus). *Fix $0 < c < \infty$ and let $p = p_n$ be as in (1.3). Then the mixing time of the* KCIP *on $\Lambda(L, 2)$ satisfies*

$$C_1 n^3 \leq \tau_{\text{mix}} \leq C_2 n^3 \log(n)^{14}$$

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

Remark 1.1. We show in [32] that the mixing time on $\Lambda(L, d)$ in dimension $d \geq 3$ satisfies $n^3 \lessapprox \tau_{\text{mix}} \lessapprox n^3 \log(n)$. We conjecture that $\tau_{\text{mix}} \approx n^3$ for $d \geq 3$ and $\tau_{\text{mix}} \approx n^3 \log(n)$ for $d = 2$.

For comparison, the mixing time of the simple random walk on $G = \Lambda(L, d)$ is known to grow like $\tau_{\text{mix}}^{\text{RW}} \approx n^{\frac{2}{d}}$ for fixed $d \in \mathbb{N}$ (see, *e.g.*, Theorem 5.5 of [21]), while the worst-case expected hitting time of 0 is given by $\tau_{\text{hit}} \approx n$ when $d \geq 3$ and $\tau_{\text{hit}} \approx n \log(n)$ when $d = 2$ (see, *e.g.*, Theorem 4 of [34]).

In the statement of Theorem 1 and throughout the paper, we assume that the quantity $0 < c < \infty$ is fixed; only $n = L^2$ grows. In particular, in Theorem 1 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. 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.

The main difficulty in extending the results of [32] to the case $d = 2$ stems from the fact that the mixing time of simple random walk on the torus is very small compared to the size of the torus in dimensions $d \geq 3$, while this is no longer the case in dimension $d = 2$. As a consequence of this fact, the behaviour of the FA1f diverges substantially from the behaviour of coalescing random walks long before all the walkers have coalesced (see [26,34]). Thus, in dimension $d = 2$, we can no longer rely on comparing the KCIP directly to the coalescing process until the number of particles is close to equilibrium, which was the main technique of [32]. Instead, we now need to analyze the behaviour of the process when it is moderately far from equilibrium. Although we focus on the special case of the torus in these papers, we believe that these behaviours are typical of the KCIP on rapidly-mixing and slowly-mixing graphs respectively.

1.1. *Related work*

KCIP models have attracted a great deal of interest recently, including applications to combinatorics, computer science, and other areas. The recent survey [15] discusses KCIPs throughout physics, while [7,8] have useful surveys of places that the KCIP has appeared outside of the physics literature. The paper [3] studies the FA1f model in one dimension. Recent mathematical progress has included new bounds on the mixing properties of KCIPs in various regimes [4–8,20,23], and the very recent work [24] makes substantial progress towards a "universal" approach for bounding relaxation times of kinetically-constrained processes. We also mention our recent related analysis [9] of the square plaquette model; this is a more complicated stochastic process that is intended to have similar qualitative behaviour to the FA1f model with more physically realistic dynamics.

1.2. *General notation*

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 $\limsup_{n\to\infty} \frac{|y(n)|}{|x(n)|} \leq C < \infty$ and $y = o(x)$ for $\limsup_{n\to\infty} \frac{|y(n)|}{|x(n)|} = 0$. We write $y = \Theta(x)$ if both $y = O(x)$ and $x = O(y)$. Finally, we also write $y \lessapprox x$ or $x \gtrapprox y$ for $y = O(x)$, and $y \approx x$ for $y = \Theta(x)$, during calculations.

2. A roadmap for the proof

Our proof strategy builds on and improves the approach [32]. We recall some notation from that paper, give a sketch of our proof of Theorem 1, and then explain where our refinements occur.

First, we note that there is an obvious bijection between the points of $\Omega = \{0, 1\}^G$ and the collection of sets $\tilde{\Omega} = \{S \subset G\}$: map $X \in \tilde{\Omega}$ to $\mathbf{1}_X \in \Omega$. We often use this bijection without explicit discusssion if there is no possibility

of confusion. For example, if $X, Y \in \Omega$, we would write $X \cap Y$ as shorthand for $1_{\{u:X[u]=Y[u]=1\}}$ or $|X|$ as shorthand for $\sum_u X[u]$.

For $1 \leq k \leq \frac{n}{2}$, let $\Omega_k \subset \Omega$ be configurations of k particles for which no two particles are adjacent, *i.e.*,

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

Also set $\Omega' = \Omega \setminus \bigcup_{k=1}^{\frac{n}{2}} \Omega_k$. For each $k \leq \frac{n}{2}$, we will denote by $\tau_{\text{mix}}^{(k)}$ the mixing time of the trace of X_t on Ω_k (see Definition 4.1 of Section 4 for the precise definition of the trace of a Markov chain). We denote by $\tau_{\text{mix}}^{(\leq k)}$ the mixing time of the trace of X_t on $\bigcup_{i \leq k} \Omega_i$. Define the quantity

$$\text{Occ}_k(\varepsilon, N) = \sup_{x \in \Omega} \inf \left\{ T \geq 1 : X_1 = x, \mathbb{P}\left(\sum_{s=1}^{T} 1_{X_s \in \bigcup_{i \leq k} \Omega_i} > N \right) > 1 - \varepsilon \right\}.$$

For a fixed N and small ε, $\text{Occ}_k(\varepsilon, N)$ denotes the first time at which the occupation measure of X_t in $\bigcup_{i \leq k} \Omega_i$ exceeds N with probability at least $(1 - \varepsilon)$.

Our proof strategy for the upper bound in Theorem 1 entails the following steps:

Step 1. We show that for a universal constant $r = r(c)$ depending only on the constant c from (1.3), and slowly-growing sequence $k_{\max} = k_{\max}(c, n) \equiv r(c) \log(n)$,

$$\tau_{\text{mix}} = O\left(\tau_{\text{mix}}^{(\leq k_{\max})} + \text{Occ}_{k_{\max}}\left(\frac{1}{8k_{\max}}, C\tau_{\text{mix}}^{(\leq k_{\max})} \right) \right).$$

This is an immediate consequence of Lemma 2.1 of [31].

Step 2. By a comparison argument using the simple exclusion process, we show that

$$\tau_{\text{mix}}^{(k)} = O\left(n^3 \log(n)^3 \right) \tag{2.2}$$

uniformly in $1 \leq k \leq k_{\max}$. See Lemma 4.2.

Step 3. By coupling the KCIP to a 'colored' version of the coalesence 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'

$$\mathbb{E}[V_{t+\varepsilon S(n)} - V_t | X_t] \leq -\delta V_t + C(n) \tag{2.4}$$

for some characteristic time scale $S(n) \approx n^3$ and bias size $C(n) \approx \log(n)$, and for fixed $\varepsilon, \delta > 0$ independent of n. See Theorem 3.1.

Step 4. By another comparison argument, we show that

$$\tau_{\text{mix}}^{(\leq k_{\max})} = O\left(\max_{1 \leq k \leq k_{\max}} \tau_{\text{mix}}^{(k)} \log(n)^{13} \right).$$

See Lemma 5.1.

Step 5. Conclude from *Step* 3 and *Step* 4 that $\text{Occ}_k(\frac{1}{8k_{\max}}, C\tau_{\text{mix}}^{(\leq k_{\max})}) = O(n^3 \log(n)^{13})$. See Proposition 6.1.

The result then follows immediately by combining the bounds in *Steps* 1, 4 and 5.

This is quite similar to the road map in [32]. The key difference occurs at *Step* 3. In [32], Inequality (2.4) was proved directly when $d \geq 3$ with $S(n) = n^3$ and $C(n) = C < \infty$ constant. The analogous bound is false in dimension $d = 2$ for small $\varepsilon > 0$, and we instead show that it holds for $S(n) = n^3$ and $C(n) = \log(n)$ when $d = 2$. This change

means that we require stronger bounds in several of the remaining stages of the proof. The version of Inequality (2.4) in this paper establishes that $V_t \lessapprox \log(n)$ with large probability after an initial burn-in period of length $T \lessapprox n^3 \log(n)$. This is much weaker than the bound $V_t \lessapprox 1$ obtained in [32], and so we now need the comparison bounds in *Step* 2 and *Step* 4 above to hold up to $k \approx \log(n)$, rather than up to $k \approx 1$.

3. Mixing at very high density: Drift condition for V_t

Recall the process $V_t = \sum_{v \in \Lambda(L,2)} X_t[v]$ from Equation (2.3). In this section, we show roughly that $V_t = O(\log(n))$ with high probability for any $t \gg n^3 \log(n)$. The proof of this fact follows almost immediately from our proof of the analogous fact in our previous paper [32], and so we state only the small adjustments that are required. As in [32], the basic idea is to count "collisions" in a closely-related Markov chain called the *coalescent process*, then show that on average V_t decreases at a rate similar to this "collision" rate.

Define $G_t = (V(G_t), E(G_t))$ to be the induced subgraph of $\Lambda(L, 2)$ with vertices $V(G_t) = \{u \in \Lambda(L, 2) : X_t[u] = 1\}$, and define

$$\text{ConnComp}(G_t) = \text{The number of connected components of } G_t. \tag{3.1}$$

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 a drift condition on $\{V_t\}_{t \in \mathbb{N}}$, which follows almost immediately from bounds in [32]:

Theorem 3.1. *There exists some constant $0 < \varepsilon_0 = \varepsilon_0(c)$ independent of n so that for all $0 < \varepsilon < \varepsilon_0$, there exist constants $C_G = C_G(\varepsilon, c) < \infty$, $\alpha = \alpha(\varepsilon, c) > 0$ and $N = N(\varepsilon, c)$ so that, for all $n > N$,*

$$\mathbb{E}[V_{\varepsilon n^3} | V_1] \le (1 - \alpha) V_1 + C_G \log(n). \tag{3.2}$$

Before giving the proof, we recall the definition of the coalescent process on a finite graph ([10,17]):

Definition 3.2 (Coalescent process). Fix a regular graph $G = (V, E)$ 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 $O_s = \{v \in G : \exists 1 \le i \le k \text{ such that } Z_s[i] = v\}$ 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 $u_s \sim \text{Unif}(\{v \in G \setminus \{v_s\} : (v, v_s) \in E\})$ and set $I_s = \{i : Z_s[i] = v_s\}$. If $u_s \le q|O_s|$, then set $Z_{s+1}[j] = u_s$ for all $j \in I_s$ 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.

Proof of Theorem 3.1. Let $\{Z_t\}_{t \in \mathbb{N}}$ be a coalescent process on $\Lambda(L, 2)$ with V_1 initial particles. Let $L_t = |O_t|$ be the number of occupied sites of Z_t, so that $L_1 = V_1$. Inequality (4.1) of [34] states that there exists a constant $0 < C < \infty$ so that, for all $t \in \mathbb{N}$,

$$\mathbb{E}[L_t] \le Cn \frac{\log(t)}{t - 1}$$

uniformly in the number $L_1 = V_1$ of initial particles. In particular, we have

$$\mathbb{E}[L_{\varepsilon n}] \le C\varepsilon \log(n). \tag{3.3}$$

Define the *number of collisions by time s* to be

$$\mathcal{C}_s = \left|\left\{1 \le u < s : \text{ConnComp}(G_{u+1}) < \text{ConnComp}(G_u)\right\}\right|. \tag{3.4}$$

We obtain a lower bound on the number of collisions by following exactly the argument given for a similar bound in Lemma 6.15 of [32], making and propagating two minor changes:

(1) We replace Inequality (6.47) of [32] and all references to the associated Theorem 5 of [34] with our Inequality (3.3) and references to Inequality (4.1) of [34].[1]

[1] Because of the different notation, Inequality (6.47) of [32] looks slightly different from our Inequality (3.3) at first glance. In the notation of [32], our Inequality (3.3) would be written as $\mathbb{E}[\sum_i \mathbf{1}_{\mathcal{A}_2^{(i)}}] \le C \log(n)$.

(2) We replace the universal constant C first defined in Inequality (6.47) of [32] with $C\log(n)$.

For any fixed $\varepsilon > 0$, the resulting lower bound on the number of collisions is

$$\mathbb{E}[\mathcal{C}_{\varepsilon n^3}] \geq \alpha V_1 - C\log(n) \tag{3.5}$$

for some constants $0 < \alpha < 1$, $0 \leq C < \infty$ that may depend on c and ε, but which do not depend on n.

Inequality (3.2) follows by an argument identical to the proof of Theorem 6.1 of [32], with one change: we replace all references to Lemma 6.15 of [32] with references to our Inequality (3.5). The proof of Theorem 6.1 in [32] is fairly long, so we include a basic sketch of the argument here. The main idea is to couple the KCIP to a coalescence process in such a way that a positive percentage of collisions in the coalescence process occur shortly before a connected component of the KCIP is removed; this allows us to connect the bound in (3.5) to the behaviour of the KCIP. The proof itself is concerned with checking that the coupling is tight enough for this transfer of information, and also checking that only a moderate number of new particles can be spawned by the KCIP over the relevant time interval. □

4. Mixing at moderate densities: Trace of KCIP on Ω_k

In this section, we bound the mixing time of the *trace* of $\{X_t\}_{t\in\mathbb{N}}$ onto the sets Ω_k defined in Equation (2.1), for all $k = O(\log(n))$. We recall the definition of the *trace* of a Markov chain:

Definition 4.1 (Trace). Fix an irreducible Markov chain $\{Z_t\}_{t\in\mathbb{N}}$ on a finite state space Θ. For a fixed subset $S \subset \Theta$, set $\eta(0) = 0$ and for $s \in \mathbb{N}$, recursively define the sequences of times

$$\begin{aligned} \eta(s) &= \inf\big\{t > \eta(s-1) : Z_t \in S\big\}, \\ \kappa(s) &= \sup\big\{u : \eta(u) \leq s\big\}. \end{aligned} \tag{4.1}$$

The quantity κ can also be written as

$$\kappa(T) = \sum_{t=1}^{T} \mathbf{1}_{Z_t \in S}. \tag{4.2}$$

Then the *trace* $\{Z_t^{(S)}\}_{t\in\mathbb{N}}$ of the Markov chain $\{Z_t\}_{t\in\mathbb{N}}$ onto the set S is given by

$$Z_t^{(S)} = Z_{\eta(t)}. \tag{4.3}$$

Fix $1 \leq k \leq \frac{n}{2}$, and let $Q_{n,k}$ be the kernel of the trace of $\{X_t\}_{t\in\mathbb{N}}$ on Ω_k. Denote by $\tau_{n,k}$ the mixing time of $Q_{n,k}$ and denote by $1 - \beta_1(Q_{n,k})$ the spectral gap of $Q_{n,k}$ (see Equation (4.6) below for a definition of spectral gap). The key result of this section is:

Lemma 4.2 (Mixing of restricted walks). *Fix $r \geq 1$. With notation as above, there exists a constant $C = C(c, r)$ that does not depend on n so that*

$$\begin{aligned} &\tau_{n,k} \leq Cn^3\log(n)^3, \\ &\frac{1}{1-\beta_1(Q_{n,k})} \leq Cn^3\log(n)^2 \end{aligned}$$

uniformly in $1 \leq k \leq r\log(n)$ for all $n > N(c, r)$ sufficiently large.

We will proceed by using comparison theory, a tool developed for comparing the mixing properties of a Markov chain of interest to those of a similar and better-understood chain (see, *e.g.*, [29] or [11] for an introduction to this

method). We prove our estimates on $Q_{n,k}$ by comparing the log-Sobolev constants of a sequence of other well-studied Markov chains. We outline this sequence of comparison bounds, with notation collected here for easy reference:

(1) Following [32], we will first compare $Q_{n,k}$ to a sped-up and restricted version of the simple exclusion process (SE) on $\Lambda(L,2)$, whose kernel is denoted Q_{MH}; see Section 4.4. The papers [10,17] give an introduction to the simple exclusion process.
(2) We will next compare the modified version of the SE process with kernel Q_{MH} to a suitably modified Bernoulli–Laplace diffusion process, whose kernel is denoted $U_{n,k}^{\mathrm{MH}}$. The original comparison paper [29] of Diaconis and Saloff-Coste compares the usual SE process to the standard Bernoulli–Laplace diffusion process. We use an argument very similar to that started in Section 3 of [29] and completed in Section 4.6 of [30]; see Section 4.3.
(3) We use direct computations and an argument from [33] to estimate the log-Sobolev constant of our modified Bernoulli–Laplace diffusion process $U_{n,k}^{\mathrm{MH}}$. See Section 4.2.

We next recall the definitions of the simple exclusion process and the Bernoulli–Laplace diffusion process, which form the basis of our kernels Q_{MH} and $U_{n,k}^{\mathrm{MH}}$:

Definition 4.3 (Simple exclusion process on $\Lambda(L,2)$). The simple exclusion process $\{Z_t\}_{t\in\mathbb{N}}$ is a Markov chain on the finite state space

$$\Omega_{n,k}^{\mathrm{SE}} \equiv \left\{ Z \in \{0,1\}^n : \sum_i Z[i] = k \right\}. \tag{4.4}$$

To update Z_t, choose two adjacent vertices $u_t, v_t \in \Lambda(L,2)$ 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 $w \notin \{u_t, v_t\}$. We denote by $Q_{n,k}^{\mathrm{SE}}$ the associated transition kernel.

Definition 4.4 (Bernoulli–Laplace Diffusion Process). The Bernoulli–Laplace diffusion process $\{Z_t\}_{t\in\mathbb{N}}$ is a Markov chain on the finite state space $\Omega_{n,k}^{\mathrm{SE}}$ given in Equation (4.4). To update Z_t, sample

$$u_t \sim \mathrm{Unif}\big(\{i : Z_t[i] = 1\}\big),$$
$$v_t \sim \mathrm{Unif}\big(\{i : Z_t[i] = 0\}\big)$$

and set

$$Z_{t+1}[u_t] = 0,$$
$$Z_{t+1}[v_t] = 1,$$
$$Z_{t+1}[w] = Z_t[w], \quad w \notin \{u_t, v_t\}.$$

We denote by $U'_{n,k}$ the associated transition kernel and let $U_{n,k} = \frac{1}{2}U'_{n,k} + \frac{1}{2}\mathrm{Id}$.

4.1. *Comparison of Markov chains using Dirichlet forms*

Before proving the main result of this section, we recall some relevant results for comparing Dirichlet forms.

Definition 4.5 (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 as:

$$\begin{aligned}
\|f\|_{2,\pi}^2 &= \sum_{x\in X} \big|f(x)\big|^2 \pi(x),\\
V_\pi(f) &= \frac{1}{2}\sum_{x,y\in X} \big|f(x)-f(y)\big|^2 \pi(x)\pi(y),\\
\mathcal{E}_P(f,f) &= \frac{1}{2}\sum_{x,y\in X} \big|f(x)-f(y)\big|^2 P(x,y)\pi(x),\\
L_\pi(f) &= \sum_{x\in X} \big|f(x)\big|^2 \log\left(\frac{f(x)^2}{\|f\|_{2,\pi}^2}\right)\pi(x).
\end{aligned} \tag{4.5}$$

Recall that the *log-Sobolev constant* and *spectral gap* of a Markov transition matrix P are given by

$$\begin{aligned}
\alpha(P) &= \inf_{f\neq 0} \frac{\mathcal{E}_P(f,f)}{L_\pi(f)},\\
1-\beta_1(P) &= \inf_{f\neq 0} \frac{\mathcal{E}_P(f,f)}{V_\pi(f)}.
\end{aligned} \tag{4.6}$$

Let K, Q be the kernels of two $\frac{1}{2}$-lazy, aperiodic, irreducible, reversible Markov chains on a state space Θ. Assume that K has unique stationary measure μ while Q has a unique stationary measure ν.

Definition 4.6 (Paths, flows). For each $a, b \in \Theta$ with $K(x,y) > 0$, 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 and let $\Gamma = \bigcup_{a,b}\Gamma_{a,b}$. Call a function $F : \Gamma \mapsto [0,1]$ 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 Equation (4.5). If

$$\begin{aligned}
L_\nu(f) &\le C_L L_\mu(f),\\
\mathcal{E}_K(f,f) &\le C_{\mathcal{E}}\mathcal{E}_Q(f,f)
\end{aligned}$$

for all functions $f : \Theta \mapsto \mathbb{R}$, then the variational characterization of α given in formula (4.6) implies

$$\alpha(Q) \ge \frac{1}{C_L C_{\mathcal{E}}}\alpha(K). \tag{4.7}$$

Using this notation, Theorem 2.1 of [29] may be restated as:

Theorem 2 (Comparison of Dirichlet forms for general chains). *Fix notation as above and also fix a flow F. Then for any function $f : \Theta \mapsto \mathbb{R}$,*

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

where

$$\mathcal{A} = \sup_{Q(q,r)>0} \frac{1}{Q(q,r)\nu(q)} \sum_{\gamma\ni(q,r)} F[\gamma]k[\gamma]K\big(i(\gamma),o(\gamma)\big)\mu\big(i(\gamma)\big).$$

Lemma 2 of [33] may be restated as:

Lemma 4.7 (Comparison of variance and log-Sobolev constants). *Let μ, ν be measures on Θ and let $\tilde{C} = \sup_{y\in\Omega} \frac{\nu(y)}{\mu(y)}$. Then for any function f on Θ,*

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

4.2. *The log-Sobolev constant of a modified Dirichlet–Laplace diffusion processes*

Let $U_{n,k}$ be as in Definition 4.4 and let $U_{n,k}^{\mathrm{MH}}$ be the Metropolis-Hastings chain with proposal distribution $U_{n,k}$ and target distribution the uniform distribution on $\Omega_{n,k} \equiv \Omega_k$; that is, let

$$U_{n,k}^{\mathrm{MH}}(x, y) = U_{n,k}(x, y)\mathbf{1}_{x,y\in\Omega_{n,k}}$$

for $x \neq y$ and $U_{n,k}^{\mathrm{MH}}(x, x) = 1 - \sum_{x\neq y} U_{n,k}^{\mathrm{MH}}(x, y)$. We define π_{MH} to be the uniform distribution on $\Omega_{n,k}$ and π_{SE} to be the uniform distribution on $\Omega_{n,k}^{\mathrm{SE}}$. Let $\mathcal{E}_{\mathrm{U,SE}}$ and $\mathcal{E}_{\mathrm{U,MH}}$ be the Dirichlet forms associated with $U_{n,k}$ and $U_{n,k}^{\mathrm{MH}}$. The main bound in this section is:

Lemma 4.8. *Fix $0 < r < \infty$. Let $\alpha(U_{n,k}^{\mathrm{MH}})$ and $1 - \beta_1(U_{n,k}^{\mathrm{MH}})$ be the log-Sobolev constant and spectral gap of $U_{n,k}^{\mathrm{MH}}$. Then there exists some constant $C = C(c, r) < \infty$ that does not depend on n so that*

$$\alpha\big(U_{n,k}^{\mathrm{MH}}\big) \ge \frac{C}{n\log(n)^3},$$
$$1 - \beta_1\big(U_{n,k}^{\mathrm{MH}}\big) \ge \frac{C}{n\log(n)^2}$$

uniformly in $1 \le k \le r\log(n)$.

Before proving this, we recall an estimate of the log-Sobolev constant of the "perfect" transition kernel $L_{n,k}^{\mathrm{MH}}$ on $\Omega_{n,k}$, defined by

$$L_{n,k}^{\mathrm{MH}}(x, y) = \frac{1}{2|\Omega_{n,k}|} + \frac{1}{2}\mathbf{1}_{x=y}.$$

We have:

Lemma 4.9 (Log-Sobolev constant of $L_{n,k}^{\mathrm{MH}}$). *Fix $0 < r < \infty$. Let $\alpha(L_{n,k}^{\mathrm{MH}})$ and $1 - \beta_1(L_{n,k}^{\mathrm{MH}})$ be the log-Sobolev constant and spectral gap of $L_{n,k}^{\mathrm{MH}}$. Then there exists some constant $0 < C = C(c, r) < \infty$ that does not depend on n so that*

$$\alpha\big(L_{n,k}^{\mathrm{MH}}\big) \ge \frac{C}{\log(n)},$$
$$1 - \beta_1\big(L_{n,k}^{\mathrm{MH}}\big) \ge C$$

uniformly in $1 \le k \le r\log(n)$, for all $n > N(r)$ sufficiently large.

Proof. This follows immediately from an application of Inequality (3.10) of [30] and the well-known fact that the spectral gap of $L_{n,k}^{\mathrm{MH}}$ is $\Theta(1)$. □

We are now ready to prove Lemma 4.8 by comparing $U_{n,k}^{\mathrm{MH}}$ to $L_{n,k}^{\mathrm{MH}}$:

Proof of Lemma 4.8. We will apply Theorem 2. To do so, we need to define the relevant paths and flows on those paths. We assume that $n > 50k$. Before giving a careful definition for our paths, we give some intuition behind their

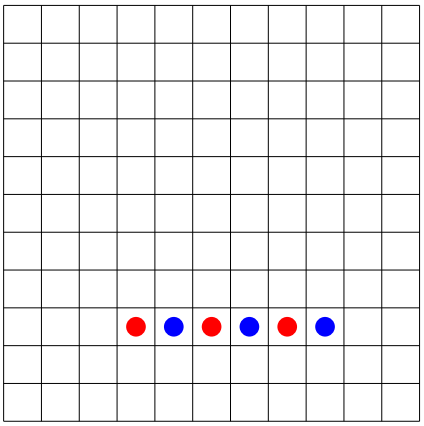

Fig. 1. The particles in X (respectively Y) are shown in red (respectively blue). Note that moving any red particle to any blue particle will result in a configuration that is not in $\Omega_{n,k}$, because the configuration will have two adjacent particles.

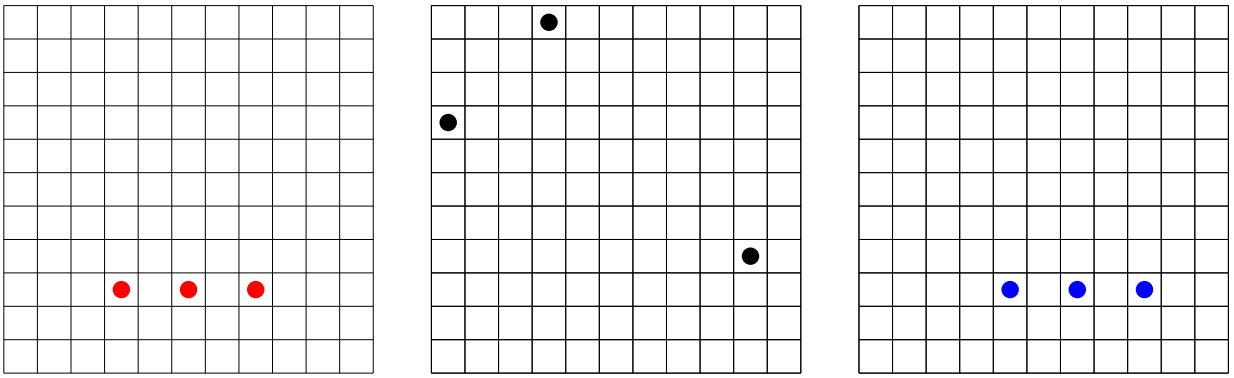

Fig. 2. The left-hand (respectively right-hand) side shows the X (respectively Y) configuration from Figure 1. The middle display shows a "typical" intermediate configuration Z. Note that the obvious paths from X to Z and from Z to Y do not exit $\Omega_{n,k}$.

construction. First, note that there is an obvious path between any pair $X, Y \in \Omega_{n,k}$: simply move particles in X to particles in Y one at a time, in any order. Unfortunately, for some choices of X, Y, this obvious path will leave the state space $\Omega_{n,k}$ – see Figure 1.

To avoid this problem, we sample a random intermediate point Z at random; with high probability, the direct paths from X to Z and from Z to Y will remain in $\Omega_{n,k}$ and the additional steps will not have a large impact on the final bound. Figure 2 shows the introduction of a random intermediate point for the bad pair of configurations in Figure 1.

Roughly speaking, since $X \cup Y \cup Z \in \Omega_{n,3k}$ with very high probability under this construction, we will almost be able to ignore the effects of the constraint that no two particles may be adjacent when doing the following calculations.

We now define the paths precisely:

Definition 4.10 (Flows for Bernoulli–Laplace diffusions). Fix $X, Y \in \Omega_{n,k}$. We sample a length-2 path from X to Y by the following algorithm:

(1) Choose Z uniformly from the set

$$\Omega^{X,Y} = \left\{ Z \in \Omega_{n,k} : \sum_{|u-v|\le 1} \big(X[u] + Y[u]\big) Z[v] = 0 \right\}$$

of configurations that have no particles next to either X or Y.

(2) Let

$$\{x_1, \dots, x_k\} = \big\{u : X[u] = 1\big\},$$
$$\{y_1, \dots, y_k\} = \big\{u : Y[u] = 1\big\},$$
$$\{z_1, \dots, z_k\} = \big\{u : Z[u] = 1\big\}$$

be the location of all particles in X, Y and Z respectively, ordered uniformly at random.

(3) Define a path $P_1^{X,Y} = (\sigma_1, \dots, \sigma_{k+1})$ from the set associated with X to the set associated with Z by

$$\sigma_i[j] = x_j, \quad i \le j,$$
$$\sigma_i[j] = z_j, \quad i > j.$$

Define a path $P_2^{X,Y} = (\eta_1, \dots, \eta_{k+1})$ from the set associated with Z to the set associated with Y by

$$\eta_i[j] = z_j, \quad i \le j,$$
$$\eta_i[j] = y_j, \quad i > j.$$

(4) Return the path $P^{X,Y} = (\mathbf{1}_{\sigma_1}, \dots, \mathbf{1}_{\sigma_{k+1}}, \mathbf{1}_{\eta_2}, \dots, \mathbf{1}_{\eta_{k+1}})$ from X to Y.

Having defined the flows, we have implicitly defined the constant $\mathcal{A}$ in Theorem 2. We must now bound that constant. Fix a pair of elements (Q, R) with $U_{n,k}^{\mathrm{MH}}(Q, R) > 0$ and $Q \neq R$. By the definition of $U_{n,k}^{\mathrm{MH}}$, we must have that $|Q \setminus R| = |R \setminus Q| = 1$. For $X, Y \in \Omega_{n,k}$, let $P^{X,Y}$ be a random path as given by Definition 4.10 and let F be the associated flow. In order to bound the weight assigned to the edge (Q, R), we note that all paths have length at most $2k$, and so

$$\sum_{X,Y \in \Omega_{n,k}} \sum_{\gamma \in \Gamma_{X,Y} : (Q,R) \in \gamma} |\gamma| F[\gamma] \le 2k \sum_{X,Y \in \Omega_{n,k}} \sum_{\ell=1}^{k} \big(\mathbb{P}\big[(Q, R) = (\sigma_\ell, \sigma_{\ell+1})\big] + \mathbb{P}\big[(Q, R) = (\eta_\ell, \eta_{\ell+1})\big] \big). \tag{4.8}$$

We note that $P_1^{X,Y}$ and $P_2^{X,Y}$ are symmetric. Thus, to bound the weight (4.8) assigned to the edge (Q, R), it is enough to bound $\mathbb{P}[(Q, R) = (\sigma_\ell, \sigma_{\ell+1})]$ for all fixed $1 \le \ell \le k$ and X, Y. To do so, we note that it is possible to sample from $P_1^{X,Y}$ using the following rejection-sampling algorithm:

(1) Choose $\hat{Z}$ uniformly from the set $\{z \in \{0, 1\}^G : \sum_{v \in G} z[v] = k\}$.

(2) Let

$$\{x_1, \dots, x_k\} = \big\{u : X[u] = 1\big\},$$
$$\{y_1, \dots, y_k\} = \big\{u : Y[u] = 1\big\},$$
$$\{\hat{z}_1, \dots, \hat{z}_k\} = \big\{u : \hat{Z}[u] = 1\big\}$$

be the location of all particles in X, Y and $\hat{Z}$ respectively, ordered uniformly at random.

(3) Define a path $P_1^{X,Y} = (\sigma_1, \dots, \sigma_{k+1})$ from the set associated with X to the set associated with $\hat{Z}$ by

$$\sigma_i[j] = x_j, \quad i \le j,$$
$$\sigma_i[j] = \hat{z}_j, \quad i > j.$$

Define the associated proposal path $\hat{\gamma} = (\mathbf{1}_{\sigma_1}, \dots, \mathbf{1}_{\sigma_{k+1}}, \mathbf{1}_{\eta_2}, \dots, \mathbf{1}_{\eta_{k+1}})$.

(4) If $\hat{Z} \in \Omega^{X,Y}$, say that we *accept* this path and return the path $\hat{\gamma}$. Otherwise, say that we *reject* this choice of $\hat{Z}$ and go back to step 1 of this algorithm.

Note that this algorithm makes sense even if X, Y are not in $\Omega_{n,k}$. We note that, for $\hat{\gamma}$ as in step 3 of the algorithm, we can compute exactly

$$\sum_{X,Y \subset \{0,1\}^G : |X|=|Y|=k} \mathbb{P}\big[(Q,R) = \big(\hat{\gamma}[\ell], \hat{\gamma}[\ell+1]\big)\big] = \binom{n}{k-1}.$$

Furthermore, for $X, Y \in \Omega_{n,k}$, it is clear that

$$\mathbb{P}[\hat{Z} \text{ is rejected}] = O\left(\frac{k}{n}\right) = O\left(\frac{\log(n)}{n}\right) = o(1).$$

Combining these two bounds, we have:

$$\sum_{X,Y \in \Omega_{n,k}} \mathbb{P}\big[(Q,R) = (\sigma_\ell, \sigma_{\ell+1})\big] \leq \binom{n}{k-1}\big(1+o(1)\big).$$

Combining this with Inequality (4.8), we have

$$\sum_{X,Y \in \Omega_{n,k}} \sum_{\gamma \in \Gamma_{X,Y} : (Q,R) \in \gamma} |\gamma| F[\gamma] \leq 2k \frac{n^{k-1}}{(k-1)!}\big(1+o(1)\big). \tag{4.9}$$

Note that $U^{\mathrm{MH}}_{n,k}$ and $L^{\mathrm{MH}}_{n,k}$ have the same stationary distribution, and that

$$\frac{U^{\mathrm{MH}}_{n,k}(x,y)}{L^{\mathrm{MH}}_{n,k}(x,y)} = \frac{n^{k-1}}{(k-1)!}\big(1+o(1)\big) \tag{4.10}$$

for any (x, y) for which $U^{\mathrm{MH}}_{n,k}(x,y) \neq 0$. Combining Inequalities (4.9) and (4.10), we conclude that our choice of flow yields a value of $\mathcal{A}$ in Theorem 2 that satisfies

$$\mathcal{A} \leq 4k^2\big(1+o(1)\big).$$

The results follow immediately from applying Theorem 2 with this bound on $\mathcal{A}$ and the bound on the log-Sobolev constant (respectively spectral gap) of $L^{\mathrm{MH}}_{n,k}$ obtained in Lemma 4.9. □

4.3. *Comparing modified Dirichlet–Laplace diffusion process to modified simple exclusion process*

For $n \in \mathbb{N}$ and $1 \leq k \leq \frac{n}{2}$, we define the graphs $G_{\mathrm{SE}} = (V_{\mathrm{SE}}, E_{\mathrm{SE}})$ and $G_{\mathrm{MH}} = (V_{\mathrm{MH}}, E_{\mathrm{MH}})$ by

$$V_{\mathrm{SE}} = \Omega^{\mathrm{SE}}_{n,k},$$
$$V_{\mathrm{MH}} = \Omega_{n,k},$$
$$E_{\mathrm{SE}} = \big\{(u,v) \in V_{\mathrm{SE}} : Q^{\mathrm{SE}}_{n,k}(u,v) > 0\big\},$$
$$E_{\mathrm{MH}} = E_{\mathrm{SE}} \cap V^2_{\mathrm{MH}},$$

where $\Omega^{\mathrm{SE}}_{n,k}$ and $Q^{\mathrm{SE}}_{n,k}$ are given in Definition 4.3, and $\Omega_{n,k} = \Omega_k$ is defined in Equation (2.1). Note that G_{MH} is a subgraph of G_{SE}.

We then define Q_{SE} to be the kernel of the $\frac{1}{2}$-lazy simple random walk on G_{SE}; this has stationary distribution π_{SE} that is uniform on V_{SE}. We define π_{MH} to be the uniform distribution on V_{MH} and define Q_{MH} to be the Metropolis-

Hastings kernel with proposal kernel Q_{SE} and stationary measure π_{MH}. That is,

$$Q_{\mathrm{MH}}(x, y) = Q_{\mathrm{SE}}(x, y)\mathbf{1}_{x,y\in V_{\mathrm{MH}}}$$

for $x \neq y$ and $Q_{\mathrm{MH}}(x, x) = 1 - \sum_{y\neq x} Q_{\mathrm{MH}}(x, y)$.

The main bound of this section is:

Lemma 4.11. *Fix $0 < r < \infty$. Let $\alpha(Q_{\mathrm{MH}})$ and $1 - \beta_1(Q_{\mathrm{MH}})$ be the log-Sobolev constant and spectral gap of Q_{MH}. Then there exists some constant $C = C(c, r) < \infty$ that does not depend on n so that*

$$\alpha(Q_{\mathrm{MH}}) \geq \frac{C}{n^2 \log(n)^3},$$

$$1 - \beta_1(Q_{\mathrm{MH}}) \geq \frac{C}{n^2 \log(n)^2}$$

uniformly in $1 \leq k \leq r \log(n)$, for all $n > N(r)$ sufficiently large.

Proof of Lemma 4.11. We will apply Theorem 2, comparing Q_{MH} to $U^{\mathrm{MH}}_{n,k}$. To do so, we need to define the relevant paths and flows on those paths. The proof of this lemma will be similar in spirit to the proof of Lemma 4.8. In both cases, we have the following intuition:

(1) there is an "obvious" direct path between pairs of points X, Y;
(2) this "obvious" path will sometimes leave the state space $\Omega_{n,k}$ of the Markov chain, and thus cannot legally be used; and
(3) we resolve this problem by choosing intermediate points according to some distribution, and then showing that the indirect path from X to Y that goes via these intermediate points will stay in $\Omega_{n,k}$ with high probability.

Unfortunately, the details of this construction are more complicated in the present lemma, so we give an expanded rough sketch of what goes wrong and how we avoid the main problems. In this sketch, we view X, Y as collections of k particles which undergo simple random walk.

We note that if $X \neq Y$ satisfy $U^{\mathrm{MH}}_{n,k}(X, Y) > 0$, then in fact $X \setminus Y$ and $Y \setminus X$ each contain only a single element, which we denote x, y respectively. The obvious path from X to Y is to send x to y along a minimal-length path, without moving any of the other particles in X and Y. Call a pair X, Y *good* if this path remains in $\Omega_{n,k}$. Since $|X|$, $|Y|$ are much smaller than n, the vast majority of choices of X, Y are good. Unfortunately, if *e.g.* the particles in X or Y are very densely packed, there may be *no* way to send x to y without exiting $\Omega_{n,k}$ or moving particles in X that already agree with Y, as illustrated in Figure 3.

Considering other configurations similar to Figure 3, it is straightforward to check that the shortest path from X to Y might involve moving at least $\sqrt{\log(n)}$ different particles in X. Thus, while *most* pairs X, Y are "good," some pairs are very bad. Moving these conflicting particles carelessly can result in very bad estimates.

To resolve this difficulty, we need to begin by moving X, Y to nearby positions X', Y' that form a "good" pair. To obtain a useful result, we have to make sure that these initial "sparsifying" moves don't have a high congestion. Since densely-packed configurations can be quite complicated, it is not obvious how to construct deterministic paths with low congestion. Fortunately, there is a simple construction of *random* paths with small congestion: allow (X, Y) to evolve according to any Markov chain whose marginal distributions have uniform distribution on $\Omega_{n,k}$; reject the resulting path if and only if the last points (X', Y') are not "good" in the above sense. As long as this rejection probability is small and the length of the path from X to X' is not too long, we will essentially be able to ignore all pairs that are not "good." See Figure 4 for sample configurations X', Y' related to the initial configurations X, Y in Figure 3.

After constructing the random paths explicitly (see Definition 4.16), almost the entire proof of this lemma is checking that the rejection probability described above is small (see Lemma 4.21). This entails checking that uniformly chosen small perturbations of *any* configuration X are unlikely to be densely packed, and that the random walk in Definition 4.16 mixes quickly enough that the endpoints of our random paths are very nearly uniformly chosen small perturbations of their starting points. We now give our formal proof.

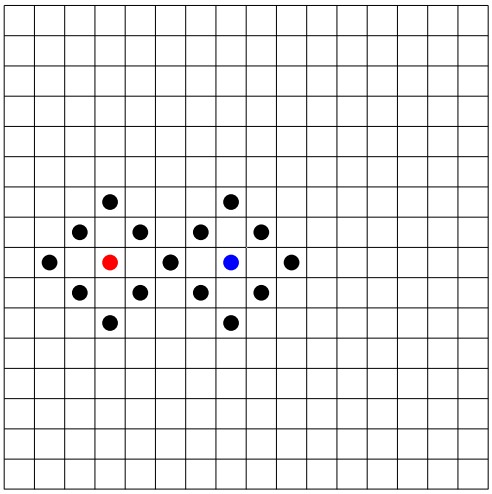

Fig. 3. The particles that are common to X and Y are shown in black. The unique particle in $X \setminus Y$ (respectively $Y \setminus X$) is shown in red (respectively blue). Note that moving the red particle to the blue particle without exiting $\Omega_{n,k}$ would require *many* other particles to move as well.

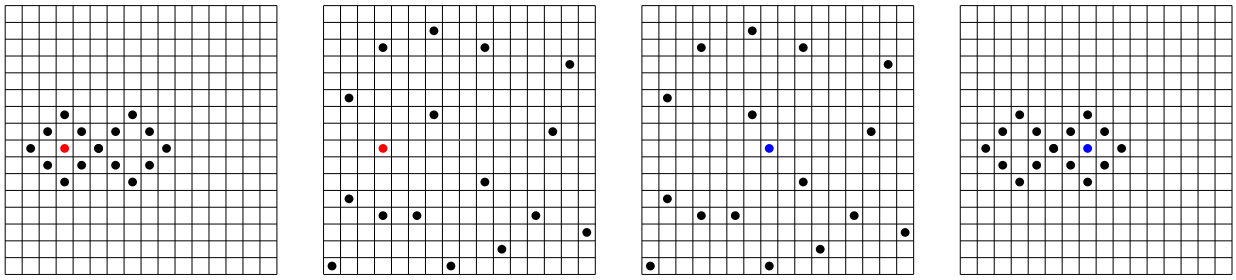

Fig. 4. The left-hand (respectively right-hand) side shows the X (respectively Y) configuration from Figure 3. The middle displays show "typical" intermediate configuration X', Y'. Note that the unique minimal-length path from X' to Y' does not exit $\Omega_{n,k}$.

Fix $X, Y \in \Omega_{n,k}$ that satisfy $U^{\mathrm{MH}}_{n,k}(X, Y) > 0$. If $X = Y$, we choose the obvious "length-0" path. This case cannot contribute to the constant $\mathcal{A}$ appearing in Theorem 2, so we ignore it for the remainder of the proof. Otherwise, $X \neq Y$ must be two configurations in $\Omega_{n,k}$ that satisfy $|\{i : X[i] = 1\} \cap \{i : Y[i] = 1\}| = k - 1$. Let x be the unique element of $\{i : X[i] = 1\} \setminus \{i : Y[i] = 1\}$ and let y be the unique element of $\{i : X[i] = 1\} \setminus \{i : Y[i] = 1\}$. As suggested above, we will construct a random path from X to Y in two steps:

(1) We construct short paths from X and Y to random configurations X' and Y' that are nearby but don't have any large clumps.
(2) We use a minimal-length path from X' to Y'.

More precisely, we write:

Definition 4.12 (Underlying paths on $\Lambda(L, 2)$). Throughout the remainder of this proof, we denote by $\Delta = \{\delta_{x,y}\}_{x,y \in \Lambda(L,2)}$ the collection of minimal-length paths between all pairs of points $x, y \in \Lambda(L, 2)$ that are described in Example 5.3 of [29]. We do not need the details of this collections of paths for our analysis - just that they are a fixed collection of minimal-length paths. These paths will be used to get between the intermediate vertices X', Y' mentioned above.

We wish to find paths that escape clumped-up figurations while remaining in $\Omega_{n,k}$. This motivates the following definition, which describes points that can easily "escape" any clump:

Definition 4.13 (Sequence of open vertices). Fix $\mathbf{1}_X = \mathbf{1}_{\{x_1,\dots,x_k\}} \in \Omega_{n,k}$, a *privileged point* $x' \in \Lambda(L,2)$ and set C. Say that a vertex $x_i \in X$ is *open* if there is a path from x_i to the outside of C that doesn't conflict with $X \cup \{x'\} \setminus \{x_i\}$ - that is, if there exists a sequence $y_1, \dots, y_\ell \in \Lambda(L,2)$ with

(1) $y_1 = x_i$ and $y_\ell \notin C$, and
(2) $|y_{j+1} - y_j| = 1$ for all $j \in \{1, 2, \dots, \ell - 1\}$, and
(3) $\min_{x \in X \cup \{x'\} \setminus \{x_i\}, 1 \le j \le \ell} |x - y_j| > 1$.

With notation as above, we say that an ordering $\sigma \in S_k$ is *a sequence of open vertices* if, for all $i \in \{1, 2, \dots, k-1\}$, $x_{\sigma[i]}$ is *open* with respect to the configuration $X \setminus \bigcup_{j<i} \{x_{\sigma[j]}\}$ and the same privileged points x' and set C.

We are particularly interested in sets that cover all "clumps" in a configuration. For fixed odd integer $m \in \mathbb{N}$ and point $x \in \Lambda(L,2)$, we denote by $C(x) = \{u : \|u - x\|_\infty \le \frac{m-1}{2}\}$ the rectangle of side length exactly m centered at x. By a small abuse of notation, for $\mathbf{1}_X \in \Omega_{n,k}$ we write $C(X) = \bigcup_{x \in X} C(x)$ and call this the size-m covering of X. We show that these sets are always associated with at least one sequence of open vertices, and furthermore that any configuration X that has many well-separated clumps will have many sequences of open vertices:

Lemma 4.14 (Existence of sequence of open vertices). *Fix $m, k \in \mathbb{N}$. For any $n > N(k,m) \equiv 100k^4m^2$ sufficiently large and any set $X = \{x_1, \dots, x_k\}$ with $\mathbf{1}_X \in \Omega_{n,k}$, privileged point $x' \in \Lambda(L,2)$ with $\min_{x \in X} |x - x'| > 1$, and set $C = C(X)$, there is an associated sequence of open vertices σ.*

Furthermore, define the graph $G = (V, E)$ with vertex set $G = X$ and edge set E given by

$$\big\{(u,v) \in E\big\} \quad \Leftrightarrow \quad \big\{\|u - v\|_1 \le 10\big\}. \tag{4.11}$$

Let $A_1, \dots, A_\ell$ be the connected components of this graph. Then any permutation σ' that preserves the same ordering as σ on each connected component (that is, which satisfies $\sigma|_{A_j} = \sigma'|_{A_j}$ for all $1 \le j \le \ell$) is also an open sequence.

Proof. In this proof, we view the torus $\Lambda(L,2)$ as a subset of the nearest-neighbour graph on the plane lattice $\mathbb{Z}^2$ using the obvious identification of points; this lets us talk about the "boundaries" of the torus in a natural way (*e.g.* the "left-most boundary" of $\Lambda(L,2)$ is the set $\{(1,y) : y \in \{1,2,\dots,L\}\}$). By the translational invariance of the torus $\Lambda(L,2)$, we can WLOG shift our embedding of $\Lambda(L,2)$ into $\mathbb{Z}^2$ so that no element of X is within distance $2m$ of any of the four boundaries of $\Lambda(L,2)$, as long as $n > 10mk$ is sufficiently large. We also note that, for $n > 100k^4m^2$, the set C does not cover any full line in $\Lambda(L,2)$.

We now prove the first claim by induction on k. For $k = 1$, it is clear that this holds for any $n \ge 1$. Thus, it is sufficient to check that there always exists at least vertex. Define

$$M^+ = \max_{1 \le i \le k} x_i[1],$$
$$M^- = \min_{1 \le i \le k} x_i[1],$$
$$z^\pm \in \big\{z \in \{x_1, \dots, x_k\} : z[1] = M^\pm\big\},$$
$$\gamma_j^\pm = \big(z^\pm[1] \pm (j-1), z^\pm[2]\big).$$

Note that $\gamma^+ = \{\gamma_j^+\}_{j \ge 1}$, $\gamma^- = \{\gamma_j^-\}_{j \ge 1}$ define two horizontal paths - the first going straight right from the right-most vertex, the second going straight left from the left-most vertex. We claim that at least one of these paths must satisfy the requirement in the definition of an open vertex. To see this, note that by maximality of the starting point M^+, γ^+ will not go within distance 1 of any point in $X \setminus \gamma^+$ before hitting the right-most boundary of $\Lambda(L,2)$. Similarly, γ^- will not go within distance 1 of any point of $X \setminus \gamma^-$ before hitting the left-most boundary of $\Lambda(L,2)$.

Thus, both paths can only be "blocked" from hitting a boundary by the single privileged point x'. Since the paths go in opposite directions, it is clear that at least one of them is moving away from x' until it gets to the boundary of $\Lambda(L,2)$. Thus, we have shown that X always contains a vertex with an allowed path to the boundary of $\Lambda(L,2)$. By the assumption that no vertex of X is within m of the boundary of $\Lambda(L,2)$, this means that X contains an open vertex. This completes the inductive argument and thus the proof of the first claim in the lemma.

To see the second claim, note that if $x \in A_j$ is an open vertex for the configuration $\mathbf{1}_{A_j}$, it is in fact an open vertex for the full configuration $\mathbf{1}_X$. Thus, to check that σ is an open sequence, it is enough to check that the restriction of σ to A_j is an open sequence for all $j \in \{1, 2, \ldots, \ell\}$. □

The following definition is useful for describing configurations that don't contain any bad "clumps" or other types of interference:

Definition 4.15 (Non-interfering locations). For configurations $X, Y \subset \Lambda(L, 2)$, set $C \supset X, Y$ and points $x, y \in \Lambda(L, 2)$, we define the collection of *non-interfering locations* as

$$\mathcal{N}(X, Y, x, y) = \left\{ u \in \Lambda(L, 2) : \min_{v \in X \cup Y \cup C^c \cup \delta_{x,y}} |u - v| > 3 \right\},$$

where $\delta_{x,y}$ is as in Definition 4.12.

We will now construct our first path segment. Our definition makes sense for any fixed r, all $n > N(r)$ sufficiently large (but not necessarily $1 \leq n \leq N(r)$ small), and all $1 \leq k \leq r \log(n)$; this is enough for our purposes. We set the notation used in the definition. Fix a parameter $T \in \mathbb{N}$, write $X = \{x_1, \ldots, x_k\}$, $Y = \{y_1, \ldots, y_k\}$ and assume $|X \cap Y| = k - 1$. Let p, q be the unique elements of $X \setminus Y$, $Y \setminus X$ respectively, and assume WLOG that $p = x_k$, $q = y_k$. Let $m = \lceil 10r^4 \log(n)^4 \rceil$, let $C = C(X \cup Y)$ be the associated covering of $X \cup Y$, and choose σ^x uniformly at random from amongst all sequences of open vertices associated with the set $X \cap Y$, covering C, and privileged point p; by Lemma 4.14, there is always at least one such sequence for all $n > N(r)$ sufficiently large. Next, choose σ^y uniformly at random from all sequences of open vertices associated with the set $X \cap Y$, covering C, and privileged point q with the *additional* property that $y_{\sigma^y[i]} \in C(x_{\sigma^x[i]})$ for all i.[2]

Definition 4.16 (First path segment: Removing clumps). We define a pair of measures $F_1^{X,Y}$ and $F_1^{Y,X}$ on paths started from $\mathbf{1}_X$ and $\mathbf{1}_Y$ respectively. We will not construct these two marginal distributions individually; instead, we define a joint distribution on paths (P_X, P_Y) with $P_X \sim F_1^{X,Y}$ and $P_Y \sim F_1^{Y,X}$. The following algorithm builds up its paths over $k - 1$ distinct stages:

(1) Set $X^{(1)} = X \setminus \{x_{\sigma^x[1]}\}$, $Y^{(1)} = Y \setminus \{y_{\sigma^y[1]}\}$.
(2) For $i \in \{1, 2, \ldots, k - 1\}$,
   (a) Let $\{Z_t^{(i)}\}_{t=1}^T$, $\{\hat{Z}_t^{(i)}\}_{t=1}^T$ be Metropolis-Hastings chains, with proposal given by $\frac{1}{2}$-lazy simple random walk on $\Lambda(L, 2)$ and target distributions being uniform on $C(x_{\sigma^x[i]}) \setminus \{u : \min_{v \in X^{(i)}} |u - v| \leq 1\}$ and $C(x_{\sigma^x[i]}) \setminus \{u : \min_{v \in Y^{(i)}} |u - v| \leq 1\}$ respectively. [3] Let the initial points of these chains be $Z_1^{(i)} = x_{\sigma^x[i]}$ and $\hat{Z}_1^{(i)} = y_{\sigma^y[i]}$ respectively. Couple these two chains so as to maximize $\mathbb{P}[Z_T^{(i)} = \hat{Z}_T^{(i)}]$.
   (b) If $\hat{Z}_T^{(i)} = Z_T^{(i)} \in \mathcal{N}(X^{(i)}, Y^{(i)}, p, q)$, define the i'th part of the path by setting $\gamma_X(i)' = (X^{(i)} \cup Z_1^{(i)}, X^{(i)} \cup Z_2^{(i)}, \ldots, X^{(i)} \cup Z_T^{(i)})$ and $\gamma_Y(i)' = (Y^{(i)} \cup \hat{Z}_1^{(i)}, Y^{(i)} \cup \hat{Z}_2^{(i)}, \ldots, Y^{(i)} \cup \hat{Z}_T^{(i)})$, letting $X^{(i+1)} = X^{(i)} \cup \{Z_T^{(i)}\} \setminus \{x_{\sigma^x[i+1]}\}$ and $Y^{(i+1)} = Y^{(i)} \cup \{\hat{Z}_T^{(i)}\} \setminus \{y_{\sigma^y[i+1]}\}$, and letting $\gamma_X(i)$, $\gamma_Y(i)$ be obtained by removing repeated elements of $\gamma_X(i)'$, $\gamma_Y(i)'$. Otherwise, say that step i *failed* and return to step (2.a).
(3) Return the paths $(\gamma_X(1), \gamma_X(2), \ldots, \gamma_X(k-1))$ and $(\gamma_Y(1), \gamma_Y(2), \ldots, \gamma_Y(k-1))$.

We denote by X' and Y' the random endpoints of these paths.

Definition 4.17 (Second path segment: Matching elements). We define a flow $F_2^{X,Y}$. Fix X, Y satisfying $|X| = |Y| = k$ and $|X \cap Y| = k - 1$. Let p, q be the unique elements of $X \setminus Y$ and $Y \setminus X$. Let $\delta_{p,q} = (z_1, \ldots, z_m)$ be as in Definition 4.12. For $1 \leq i \leq m$, define $Z_i = (X \cap Y) \cup \{z_i\}$. Define $\gamma_{X,Y} = (\mathbf{1}_{Z_1}, \ldots, \mathbf{1}_{Z_m})$. When every element of $\gamma_{X,Y}$ is an element of $\Omega_{n,k}$, $F_2^{X,Y}$ assigns weight 1 to $\gamma_{X,Y}$. Otherwise, we do not define $F_2^{X,Y}$.

[2] To see that a pair $\sigma^{(x)}, \sigma^{(y)}$ with this additional property exists, observe that for n large, the side-length m of the square $C(x_{\sigma^x[i]})$ is much larger than the largest possible size of a connected component of the graph defined in Equation (4.11). Thus the existence of such a sequence follows from the second clause of Lemma 4.14.

[3] Note that both sets are based on $C(x_{\sigma^x[i]})$; the second instance should not be replaced by $C(y_{\sigma^y[i]})$

Finally, we define a measure F on $\Gamma_{X,Y}$ by giving an algorithm for sampling from F:

Definition 4.18 (Full path). Fix a parameter $T \in \mathbb{N}$. To sample from F, run the following random algorithm:

(1) Sample paths $P_X \sim F_1^{X,Y}$, $P_Y \sim F_1^{Y,X}$ with endpoints X', Y' according to the coupling in Definition 4.16.
(2) Sample a path $P_{X',Y'} \sim F_2^{X',Y'}$. When this path is not defined, say that *the long path fails* and return to Step (1).
(3) Return the random path $(P_X, P_{X',Y'}, P_Y^{\dagger})$, where $P_Y^{\dagger}$ denotes reversing the order of a path.

Having defined the flows, we have implicitly defined the constant $\mathcal{A}$ in Theorem 2. We must now bound that constant. To do so, we consider a fixed in E_{MH} and bound the total weight of all paths that cross through the edge. Since the constant $\mathcal{A}$ is defined as a sum over all flows, we can bound the contributions due to the first type of path (see Definition 4.16) and the second type of path (see Definition 4.16) separately.

We begin by bounding the flow due to the first type of path. First, we show that with probability $1 - o(1)$, none of the $k-1$ steps in the construction of $F_1^{X,Y}$ will fail, and also the long path obtained following the initial sampling from $F_1^{X,Y}$ will not fail. Checking that, with overwhelming probability, *none* of the events fail will allow us to essentially ignore the rejection steps when estimating the weight given to any edge, at the cost of a small multiplicative constant. This substantially simplifies our analysis.

Lemma 4.19 (Local failures are rare). *Following the notation of Definition 4.16, there exists $N_0 = N_0(r)$ so that the probability $\mathbb{P}_i$ that step i fails is bounded by*

$$\mathbb{P}_i \leq \frac{1}{2(r\log(n))^{1.5}}$$

for all $T > 2\log(n)^{20}\log(\log(n))$ and all $n > N_0$ sufficiently large.

Proof. This relies on two estimates: first, showing that the random walk in Definition 4.16 mixes within T steps, and second, showing that its stationary measure always places very high weight on states that don't trigger *failure*.

Let $\{Z_t\}_{t\geq 1} = \{Z_t^{(i)}\}_{t\geq 1}$ be as in stage i of Definition 4.16, and let $\mathcal{D}$ be the connected component of its state space, $C(Z_1) \setminus \{u : \min_{v\in X^{(i)}} |u - v| \leq 1\}$, that contains Z_1. The critical estimate is the following mixing bound:

Proposition 4.20. *The mixing time τ_{mix} of $\{Z_t\}_{t\geq 1}$ on $\mathcal{D}$ satisfies*

$$\tau_{\mathrm{mix}} \leq C_1 \log(n)^{16}\log\big(\log(n)\big) \tag{4.12}$$

for some $0 < C_1 = C_1(r) < \infty$.

Proof. We note that $\mathcal{D} \subset C(Z_1)$, so

$$|\mathcal{D}| \leq \big|C(Z_1)\big| = O\big(r^8\log(n)^8\big).$$

Furthermore, all non-zero transition probabilities for $\{Z_t\}_{t\geq 1}$ are at least $\frac{1}{4}$. By Theorem 1 of [2], we have

$$\tau_{\mathrm{mix}} \leq 64|\mathcal{D}|^2\big(\log\big(|\mathcal{D}|\big) + \log(4)\big) = O\big(r^{16}\log(n)^{16}\log\big(r\log(n)\big)\big). \qquad \square$$

Having proved Proposition 4.20, we now continue with the proof of Lemma 4.19. Let the good set $\mathcal{N} = \mathcal{N}(X^{(i)}, Y^{(i)}, p, q)$ be as in stage i of Definition 4.16.

We recall that $\mathcal{D}$ is obtained from the square $C(Z_1)$ by removing $O(|X \cup Y|)$ points and then keeping the connected component containing Z_1. Since Z_1 is an open vertex, the connected component containing Z_1 must contain the entire boundary of $C(Z_1)$. Thus, by the isoperimetric inequality for squares (see *e.g.* Theorem 1.2 of [16]), the set $\mathcal{D}$ must be of size at least

$$|\mathcal{D}| \geq \big|C(Z_1)\big| - O\big(|X \cup Y|^2\big) = \big\lceil 10r^4\log(n)^4\big\rceil^2 - O\big(r^2\log(n)^2\big) = \big|C(Z_1)\big|\big(1 - O\big(r^{-6}\log(n)^{-6}\big)\big).$$

Denoting by $\partial A = \{x \in A : \min_{y \notin A} |x - y| = 1\}$ the boundary of a set $A \subset \Lambda(L, 2)$, we also have the trivial bound

$$\begin{aligned}\big|C(Z_1) \setminus \mathcal{N}\big| &= O\big(\big|\partial C(Z_1)\big| + |X \cup Y| + |\delta_{p,q}|\big) \\ &= O\big(r^4 \log(n)^4 + r\log(n) + r^4\log(n)^4\big) = O\big(r^4\log(n)^4\big).\end{aligned}$$

Putting together these two displayed equations,

$$\begin{aligned}\frac{|\mathcal{N} \cap \mathcal{D}|}{|\mathcal{D}|} &\geq \frac{|\mathcal{D}| - |C(Z_1) \setminus \mathcal{N}|}{|\mathcal{D}|} \\ &= 1 - O\left(\frac{1}{r^4 \log(n)^4}\right). \end{aligned} \tag{4.13}$$

Combining Inequalities (4.12) and (4.13), there exists some constant $A = A(r)$ so that for $T > A\log(n)^{19} \times \log(\log(n))$ and n sufficiently large,

$$\begin{aligned}\mathbb{P}[Z_T \in \mathcal{N}] &\geq \frac{|\mathcal{N} \cap \mathcal{D}|}{|\mathcal{D}|} - 2\big(2^{-\lfloor \frac{T}{\tau_{\text{mix}}} \rfloor + 1}\big) \\ &\geq 1 - O\left(\frac{1}{(r\log(n))^{1.5}}\right),\end{aligned}$$

completing the proof. □

The same argument (with easier estimates) bounds the probability of rejecting a full path:

Lemma 4.21 (Long paths rarely fail). *Following the notation of Definition 4.18, for the constants T, r as in Lemma 4.19, we have*

$$\mathbb{P}[\textit{the long path fails}] = o(1)$$

as n goes to infinity.

Proof. Let X, Y and p, q and X', Y' be as in Definition 4.18, and define $\mathcal{N} = \mathcal{N}(X^{(k-1)}, Y^{(k-1)}, p, q)$. Using the notation of Definition 4.18, we note that the path $\delta_{p,q} = (z_1, \ldots, z_m)$ depends only on the two points $p, q \in X \Delta Y$, not any further randomization. Furthermore, $\delta_{p,q}$ is a minimal-length path between p and q, and thus its intersection $|\delta_{p,q} \cap \mathcal{N} \cap \mathcal{C}|$ with the roughly-square set $\mathcal{C} \cap \mathcal{N}$ is of size $O(\sqrt{|\mathcal{N} \cap \mathcal{C}|}) = O(r^4 \log(n)^4)$. Therefore, by the same calculation used to obtain Inequality (4.13), we have

$$\frac{|\delta_{p,q} \cap \mathcal{N} \cap \mathcal{C}|}{|\mathcal{N} \cap \mathcal{C}|} = o(1).$$

Combining this with Inequality (4.12) completes the proof. □

Next, we show that this implies the total weight given to any particular edge is small. The main idea is that it would be possible to do an *exact* calculation of the weight given to any edge if the rejection probability were exactly 0; since the rejection probability is close to 0, this exact computation gives a very good bound on the weight:

Lemma 4.22 (Contribution of first path type). *Fix $r > 0$ and $1 \leq k \leq r\log(n)$. Following the notation of Definition 4.16 and fixing $T \geq \log(n)^{20}\log(\log(n))$ so that Lemmas 4.19 and 4.21 apply, we have for all distinct $A, B \in \Omega_{n,k}$ satisfying $Q_{\text{MH}}(A, B) > 0$ that*

$$\sum_{X,Y : |X \cap Y| = k-1} \sum_{\gamma \in \Gamma_{X,Y} : (A,B) \in \gamma} F_1^{X,Y}[\gamma] \leq 8nk^2(T+1)$$

for all $n > N_0(r)$ sufficiently large.

Proof. Fix a configuration $A \in \Omega_{n,k}$, choose $X \sim \mathrm{Unif}(\Omega_{n,k})$ and then choose $Y \sim \mathrm{Unif}(\{y \in \Omega_{n,k} : |X \cap y| = k-1\})$, and choose a random path $\gamma_X = (\gamma_X(1), \gamma_X(2), \ldots, \gamma_X(k-1))$ according to Definition 4.16. For $1 \le i \le k-1$, note that we can write $\gamma_X(i) = (X^{(i)} \cup Z_1^{(i)}, X^{(i)} \cup Z_2^{(i)}, \ldots, X^{(i)} \cup Z_T^{(i)})$ as in that definition. Using this notation,

$$\sum_{X,Y:|X\cap Y|=k-1} \sum_{\gamma\in\Gamma_{X,Y}:(A,B)\in\gamma} F_1^{X,Y}[\gamma] \le n\binom{n}{k} \sum_{i=1}^{k-1}\sum_{t=0}^{T} \mathbb{P}\big[A = X^{(i)} \cup Z_t^{(i)}\big]. \tag{4.14}$$

Thus, it is enough to bound the probabilities $\mathbb{P}[A = X^{(i)} \cup Z_t^{(i)}]$ for all $1 \le i \le k-1$ and $0 \le t \le T$. We begin by bounding these probabilities in the special case $i = 1$. Noting that all particles in X are open with probability $1 - o(1)$, we have by the usual 'birthday problem' bound that[4]

$$\mathbb{P}\big[X^{(1)} = S\big] \le \frac{2}{\binom{n}{k-1}} \tag{4.15}$$

for all $S \in \Omega_{n,k-1}$ and all $n > N_0(r)$ sufficiently large.

Next, we note that the following is a valid rejection-sampling algorithm for choosing $x_{\sigma^x[1]}$ conditional on $X^{(1)}$:

(1) Sample $\hat{x}$ uniformly at random from among all elements of the largest connected component $\mathcal{D}_1$ of $\Lambda(L,2) \setminus X^{(1)}$.
(2) Let $p = p(\hat{x}, X^{(1)})$ be the percentage of all sequences of open vertices for configuration $X^{(1)} \cup \{\hat{x}\}$ that begin with $\hat{x}$. Then *accept* $\hat{x}$ with probability p; otherwise *reject* and go back to step (1).

Before analyzing the "corrected" choice of $x_{\sigma[1]}$ conditional on $X^{(1)}$, we analyze the "uncorrected" choice of $\hat{x}$. Sample $\hat{x}$ (conditional on $X^{(1)}$) uniformly at random from among all elements of the largest connected component $\mathcal{D}_1$ of $\Lambda(L,2) \setminus X^{(1)}$. Let $\{\hat{Z}_t^{(1)}\}_{t\in\mathbb{N}}$ be the Markov chain on $\mathcal{D}_1$ constructed as in Step (2.a) of Definition 4.16, started at $\hat{Z}_1^{(1)} = \hat{x}$. Since $\hat{x}$ was drawn from the stationary measure of $\{\hat{Z}_t^{(1)}\}_{t\in\mathbb{N}}$, we have

$$\mathbb{P}\big[\hat{Z}_t^{(1)} = z\big] = \frac{1}{|\mathcal{D}_1|} \tag{4.16}$$

for all $z \in \mathcal{D}_1$ and all $t \in \mathbb{N}$. By the above rejection-sampling algorithm for $x_{\sigma[1]}$ and the obvious bounds $\frac{1}{k} \le p \le 1$, this implies that the true path $\{Z_t^{(1)}\}_{t\in\mathbb{N}}$ that appears in Step (2.a) of Definition 4.16 satisfies

$$\mathbb{P}\big[Z_t^{(1)} = z\big] \le \frac{k}{|\mathcal{D}_1|} \tag{4.17}$$

for all $z \in \mathcal{D}_1$ and $t \in \mathbb{N}$. Combining Inequalities (4.15) and (4.17), with the bound (4.13) on $|\mathcal{D}_1|$, we conclude that

$$\mathbb{P}\big[X^{(1)} \cup Z_t^{(1)} = S\big] \le \frac{4k}{\binom{n}{k}} \tag{4.18}$$

for all $S \in \Omega_{n,k}$ and all $0 \le t \le T$, whenever $n > N_0(r)$ is sufficiently large.

Analogous bounds for $1 < i \le k-1$ will follow by Proposition 4.20. In particular, let $\mathcal{D}_i$ be the largest connected component of $\Lambda(L,2) \setminus X^{(i)}$. By Proposition 4.20 and the same argument giving Inequality (4.15), we have

$$\mathbb{P}\big[X^{(i)} = S\big] \le \frac{2}{\binom{n}{k-1}} + in^{-10} \tag{4.19}$$

for all $S \in \Omega_{n,k-1}$, all $1 \le i \le k-1$ and all $n > N_0(r)$ sufficiently large. Similarly, by Proposition 4.20 and the same argument giving Inequality (4.18), we have

$$\mathbb{P}\big[X^{(i)} \cup Z_t^{(i)} = S\big] \le k\left(\frac{4}{\binom{n}{k}} + 2in^{-10}\right) \tag{4.20}$$

for all $S \in \Omega_{n,k}$, all $1 \le i \le k-1$ and all $n > N_0(r)$ sufficiently large.

[4] See *e.g.* the survey [12] for exact calculations related to the birthday problem.

Combining Inequalities (4.14) and (4.20) and applying Lemma 4.21 completes the proof of the lemma. □

The following bound on the contribution of the second path type follows immediately from Example 5.3 of [29] and Lemmas 4.19 and 4.21:

Lemma 4.23 (Contribution of second path type). *Fix $r > 0$ and $1 \le k \le r\log(n)$. For $X, Y \in \Omega_{n,k}$ satisfying $|X \cap Y| = k-1$, let the random variables $X' = X'(X,Y)$ and $Y' = Y'(X,Y)$ be as in Definition 4.16. Following the notation of Definition 4.17 and fixing $T \ge \log(n)^{20}\log(\log(n))$, we have for all distinct $A, B \in \Omega_{n,k}$ satisfying $Q_{\mathrm{MH}}(A,B) > 0$ that*

$$\sum_{X,Y:|X\cap Y|=k-1}\sum_{x',y'\in\Omega_{n,k}} \mathbb{P}\big[(X',Y') = (x',y')|X,Y\big] \sum_{\gamma\in\Gamma_{x',y'}:(A,B)\in\gamma} F_2^{x',y'}[\gamma] \le 16n^{1.5}$$

for all $n > N_0(r)$ sufficiently large.

Proof. Fix a configuration $A \in \Omega_{n,k}$ and a pair $X, Y \in \Omega_{n,k}$ satisfying $|X \cap Y| = k-1$. Let $x = x(X,Y)$ and $y = y(X,Y)$ be the unique elements of $X \setminus Y$ and $Y \setminus X$ respectively. By Lemmas 4.19 and 4.21,

$$\mathbb{P}\big[(X',Y') = S|x(X,Y), y(X,Y)\big] \le \frac{4}{\binom{n}{k}^2}$$

for all $S \subset \Omega_{n,k}^2$ and all $n > N_0(r)$ sufficiently large. In particular, the probability mass function of (X', Y') conditional on $x(X,Y)$ and $y(X,Y)$ is bounded by a constant factor times the probability mass function of the uniform distribution.

By the same calculation as in Example 5.3 of [29], this implies

$$\mathbb{P}[A \in P_{X',Y'}] \le \frac{16\sqrt{n}}{\binom{n}{k}}$$

for all $n > N_0(r)$ sufficiently large. The result follows immediately by the same bound as Inequality (4.14). □

Combining Lemmas 4.22 and 4.23, and noting that all paths have length at most $k(T+1)+2\sqrt{n}$, we have

$$\begin{aligned}\mathcal{A} &\le \left(\max_{x,y,q,r:Q_{\mathrm{MH}}(q,r)>0,x\ne y}\frac{U_{n,k}^{\mathrm{MH}}(x,y)}{Q_{\mathrm{MH}}(q,r)}\right)\times\left(\max_{\gamma:F(\gamma)>0}|\gamma|\right)\times\left(\max_{(q,r):Q_{\mathrm{MH}}(q,r)>0}\sum_{\gamma\ni(q,r)}F[\gamma]\right)\\ &\le \left(\frac{2}{n}\right)\times\big(k(T+1)+2\sqrt{n}\big)\times\big(8nk^2(T+1)+16n^{1.5}\big)\end{aligned}$$

for all $n > N(c,r)$ sufficiently large.

Lemma 4.11 now follows immediately from an application of Theorem 2, with comparison provided by Lemma 4.8. □

4.4. *Comparison of modified simple exclusion process to KCIP*

Let $\alpha(Q_{n,k}), 1-\beta_1(Q_{n,k})$ be the log-Sobolev constant and spectral gap of $Q_{n,k}$, and let $\alpha(Q_{\mathrm{MH}})$ and $1-\beta_1(Q_{\mathrm{MH}})$ be the log-Sobolev constant of Q_{MH}. As shown in Inequality (5.10) of [32], we have:

$$\begin{aligned}\alpha(Q_{n,k}) &\ge \frac{1}{4n}\alpha(Q_{\mathrm{MH}}),\\ 1-\beta_1(Q_{n,k}) &\ge \frac{1}{4n}\big(1-\beta_1(Q_{\mathrm{MH}})\big).\end{aligned} \tag{4.21}$$

4.5. *Proof of Lemma 4.2*

We put together the bounds obtained in Sections 4.2, 4.3 and 4.4:

Proof of Lemma 4.2. By Lemma 4.11 and Inequality (4.21), the log-Sobolev constant $\alpha(Q_{n,k})$ and spectral gap $1-\beta_1(Q_{n,k})$ of $Q_{n,k}$ satisfy

$$\alpha(Q_{n,k}) \geq \frac{C}{n^3 \log(n)^3},$$
$$1-\beta_1(Q_{n,k}) \geq \frac{C}{n^3 \log(n)^2}$$

for some $C = C(c,r)$ for all $n > N(c,r)$ sufficiently large. Lemma 4.2 follows immediately from an application of Inequality (3.3) of [30]. □

5. Mixing at moderate density: Main bounds

We set some notation for the remainder of this section. For fixed $1 \leq k \leq n$, let $\{Y_t\}_{t\in\mathbb{N}}$ be the trace of $\{X_t\}_{t\in\mathbb{N}}$ on $\bigcup_{i=1}^{k} \Omega_{n,i}$, let $P_{n,k}$ be the transition kernel of $\{Y_t\}_{t\in\mathbb{N}}$, and let $\tau_{\text{mix}}^{(\leq k)}$ be the mixing time of $P_{n,k}$. The main result of this section is:

Lemma 5.1 (Mixing at Moderate Density). *Fix $0 < r < \infty$. There exists a constant $C = C(r,c)$ so that*

$$\tau_{\text{mix}}^{(\leq k)} \leq C n^3 \log(n)^{13}$$

uniformly in $1 \leq k \leq r\log(n)$.

Our strategy is to use Theorem 1.1 of [22], along with some soft bounds, to 'glue together' the bounds on $\{\tau_{\text{mix}}^{(k)}\}_{1\leq k\leq r\log(n)}$ from Section 4. The basic idea of [22] (as well as recent related papers [19,31]) is that it is possible to bound the relaxation time of a Markov chain on a state space Θ decomposed as $\Theta = \bigcup_{i=1}^{m} \Theta_i$ by bounding the relaxation times of certain "restricted" chains on $\Theta_1, \ldots, \Theta_m$ and the relaxation time of a "projected" chain on $\{1,2,\ldots,m\}$ that measures the typical transition rates between $\Theta_1, \ldots, \Theta_m$ near stationarity.

In our case, we will be able to easily compare our "projected" chain to biased random walk on the path $\{1,2,\ldots,k-1\}$. The bounds on the "restricted" chains will come from combining two pieces: the results in Section 4, and very soft bounds on the transition rates between $\Omega_{n,k}$ and $\Omega_{n,k\pm1}$. These latter bounds will be obtained by coupling the KCIP to a simple exclusion process over short time intervals and using explicit calculations for the simple exclusion process.

Although explicitly giving these comparisons will require us to develop some additional notation, we emphasize that the estimates in this section all come from fairly explicit calculations for well-studied Markov chains.

5.1. *Review of results in* [22]

Fix $k \in \mathbb{N}$. Let P be the transition kernel of the KCIP $\{X_t\}_{t\in\mathbb{N}}$, and for $1 \leq i \leq k-1$, let P_i be the *restriction* of P to $\Omega_i \cup \Omega_{i+1}$, defined by:

$$P_i(x,y) = \begin{cases} P(x,y), & x \neq y, x, y \in \Omega_i \cup \Omega_{i+1}, \\ 1 - \sum_{y\in\Omega_i\cup\Omega_{i+1}} P(x,y), & x = y, \\ 0, & \text{otherwise.} \end{cases}$$

Also define the kernel $\tilde{P}$ on the discrete set $\{1,2,\ldots,k-1\}$ by

$$\tilde{P}(i,j) = \frac{\pi((\Omega_i \cup \Omega_{i+1}) \cap (\Omega_j \cup \Omega_{j+1}))}{3\pi(\Omega_i \cup \Omega_{i+1})}, \quad i \neq j,$$

$$\tilde{P}(i,i) = 1 - \sum_{j \neq i} \tilde{P}(i,j).$$

Theorem 1.1 of [22] implies:

Theorem 3. *With notation as above*,

$$1 - \beta_1(P_{n,k}) \geq \frac{1}{9}\big(1 - \beta_1(\tilde{P})\big) \min_{1 \leq i \leq k-1} \big(1 - \beta_1(P_i)\big).$$

In the remainder of this section, we bound the terms in Theorem 3.

5.2. *Bounds on P_i and $\tilde{P}$*

We obtain bounds on the spectral gaps of the kernels $\{P_i\}_{i=1}^{k-1}$ and $\tilde{P}$ defined in Section 5.1.

5.2.1. *Bound on $\tilde{P}$*
We begin by bounding $1 - \beta_1(\tilde{P})$:

Lemma 5.2. *Fix $0 < r < \infty$. Then there exists $C = C(r,c)$ so that*

$$1 - \beta_1(\tilde{P}) \geq \frac{C}{\log(n)^2}$$

uniformly in $1 \leq k \leq r\log(n)$.

Proof. We will first obtain bounds on the hitting times of certain large sets for a reversible Markov chain Z_t evolving according to $\tilde{P}$. Then we use Theorem 1.1 of [28] (see also [25]) to obtain a mixing time estimate for $\tilde{P}$ from our bound on its hitting times. Fortunately, $\tilde{P}$ is a birth-and-death chain, so explicit formulas for the relevant hitting times are available.

We can assume without loss of generality that $k = \lfloor r\log(n) \rfloor$. We begin by expanding our formula for $\tilde{P}$. For $1 \leq i < k$, the usual 'birthday problem' bound gives[5]

$$\begin{aligned}
\tilde{P}(i,i+1) &= \frac{\pi(\Omega_{i+1})}{3\pi(\Omega_i \cup \Omega_{i+1})} \\
&= \frac{1}{3} \frac{|\Omega_{i+1}|(\frac{c}{n})^{i+1}(1-\frac{c}{n})^{n-i-1}}{|\Omega_i|(\frac{c}{n})^{i}(1-\frac{c}{n})^{n-i} + |\Omega_{i+1}|(\frac{c}{n})^{i+1}(1-\frac{c}{n})^{n-i-1}} \\
&= \frac{1}{3} \frac{c}{i+1+c} \left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right).
\end{aligned}$$

Similarly, for $1 < i \leq k$,

$$\begin{aligned}
\tilde{P}(i,i-1) &= \frac{\pi(\Omega_i)}{3\pi(\Omega_i \cup \Omega_{i+1})} \\
&= \frac{1}{3} \frac{|\Omega_i|(\frac{c}{n})^{i}(1-\frac{c}{n})^{n-i}}{|\Omega_i|(\frac{c}{n})^{i}(1-\frac{c}{n})^{n-i} + |\Omega_{i+1}|(\frac{c}{n})^{i+1}(1-\frac{c}{n})^{n-i-1}} \\
&= \frac{1}{3} \frac{i+1}{i+1+c} \left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right).
\end{aligned}$$

[5] See *e.g.* the survey [12] for exact calculations related to the birthday problem.

Finally, for all $1 \le i \le k$,

$$\begin{aligned}\tilde{P}(i,i) &= 1 - \tilde{P}(i,i+1) - \tilde{P}(i,i-1)\\ &\ge \frac{1}{3}\left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right),\end{aligned}$$

where by convention $\tilde{P}(1,0) = \tilde{P}(k,k+1) = 0$. Similarly, the stationary distribution $\tilde{\mu}$ of $\tilde{P}$ satisfies

$$\begin{aligned}\tilde{\mu}(i) &= \frac{1}{Z}\binom{n}{i}\left(\frac{c}{n}\right)^i\left(1-\frac{c}{n}\right)^{n-i}\left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right),\\ Z &= \sum_{j=1}^{k}\binom{n}{i}\left(\frac{c}{n}\right)^i\left(1-\frac{c}{n}\right)^{n-i}\left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right).\end{aligned}$$

Let $m^- = \max(1, \lfloor \frac{c}{4} \rfloor)$, $m^+ = \lceil 4\max(1,c) \rceil$. For $\{Z_t\}_{t\in\mathbb{N}}$ a Markov chain with transition kernel $\tilde{P}$, let

$$\begin{aligned}\tau^- &= \min\{t \in \mathbb{N} : Z_t = m^-\},\\ \tau^+ &= \min\{t \in \mathbb{N} : Z_t = m^+\}\end{aligned}$$

be the first hitting times of m^- and m^+ respectively. By standard formulas (see *e.g.*, [27]),

$$\begin{aligned}\mathbb{E}[\tau^+ | Z_1 = 1] - 1 &= \sum_{v=1}^{m^+-1}\left(\frac{1}{\tilde{\mu}(v)\tilde{P}(v,v+1)}\sum_{q=1}^{v}\tilde{\mu}(q)\right)\\ &\le \left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right)\frac{1}{3c}\sum_{v=1}^{m^+-1}\left(\frac{v+1+c}{\binom{n}{v}(\frac{c}{n})^v}\sum_{q=1}^{v}\binom{n}{q}\left(\frac{c}{n}\right)^q\right)\\ &\le \left(1 + O\left(\frac{r^2\log(n)^2}{n}\right)\right)\frac{1}{3c}\sum_{v=1}^{m^+-1} v(v+1+c)\\ &= O\left((m^+)^3\right) = O(k^2).\end{aligned} \tag{5.1}$$

Using the same argument we can also obtain that

$$\mathbb{E}[\tau^- | Z_1 = k] = O(k^2). \tag{5.2}$$

We also note that $\sum_{v=1}^{m^+}\tilde{\mu}(v) > 0.501$, $\sum_{v=m^-}^{k}\tilde{\mu}(v) > 0.501$ for all n sufficiently large. Combining this fact with Inequalities (5.1) and (5.2), Theorem 1.1 of [28] implies that the mixing time $\tilde{\tau}_{\text{mix}}$ of $\tilde{P}$ satisfies

$$\tilde{\tau}_{\text{mix}} = O(k^2).$$

Since the mixing time of a Markov chain bounds its relaxation time, this completes the proof. □

5.2.2. *Bounds on P_i*

Next, we will bound the spectral gap of P_i. This will follow from the bounds in Section 4 on the bounds of the trace processes on Ω_i and Ω_{i+1}, combined with some very rough bounds on the transition time between Ω_i and Ω_{i+1}. The remainder of this section is devoted to computing these rough bounds. Although we give additional details in the proofs, the remainder of the arguments in Section 5.2 follow quickly from the following observations:

(1) It is straightforward to check that, whenever $X_t \in \Omega_i$ contains two particles that are within distance 3 of each other, the probability of moving from Ω_{i+1} to Ω_i within $O(n^2)$ steps is bounded away from 0.

(2) It is possible to couple the trace of $\{X_t\}_{t\in\mathbb{N}}$ on Ω_i to a $(1-\frac{1}{n})$-lazy version of the simple exclusion process $\{Y_t\}_{t\in\mathbb{N}}$ with i particles so that, with high probability, $X_t = Y_t$ until the first time that any two particles of the simple exclusion process get within distance 3. We call such a time a "near-collision time."
(3) The rate of "near-collision times" associated with the simple exclusion process are very well-understood (see *e.g.* [34] and [26]).

This means that we can bound the transition times between Ω_i and Ω_{i+1} by translating existing results on near-collision times of the simple exclusion process. The coupling mentioned in item (2) of the above sequence of observations is the obvious step-by-step maximal coupling, and so we do not give an explicit construction. Such an explicit construction is available in Section 7 of [32].

To obtain the required bounds, we first need a bound on near-collisions for the simple exclusion process. For $i \in \mathbb{N}$, let

$$\mathcal{H}(i) = \left\{ X \in \Omega_i : \min_{u,v:X[u]X[v]=1} |u-v| > \frac{\sqrt{n}}{\log(n)^{0.25}} \right\}$$

be the collection of very well-spaced configurations in Ω_i, and define

$$\mathcal{G}(i) = \left\{ Y \in \Omega_{i+1} : \exists X \in \mathcal{H}, x \in \Lambda(L,2) \text{ s.t. } Y = X \cup \{x\}, \min_{u:X[u]=1} |u-x| = 2 \right\} \tag{5.3}$$

to be the collection of all configurations in Ω_{i+1} consisting of a well-spaced configuration in Ω_i with one additional particle added near to an existing particle. We need the following bound on collision times for the simple exclusion process:

Lemma 5.3 (Hitting from well-spaced configurations). *Fix $m \in \mathbb{N}$ and $1 \le i \le m$. Let $\{S_t\}_{t\in\mathbb{N}}$ be a simple exclusion process started at a configuration $S_1 \in \mathcal{G}(i)$ and let*

$$\tau_{\text{coll}} = \min\{t > 0 : \exists u, v \in \Lambda(L,2) \text{ s.t. } S_t[u] = S_t[v] = 1 \text{ and } |u-v| = 1\} \tag{5.4}$$

be the first time that a collision occurs. Then there exists some $\delta = \delta(m) > 0$ so that

$$\mathbb{P}\left[\tau_{\text{coll}} > \delta \frac{n^2}{\log(n)} \right] > \frac{\delta}{\log(n)}.$$

Proof. Let $S = \{u : S_1[u] = 1\}$ and let x_1, x_2 be two elements of S at distance exactly 2. Existing bounds on the first collision time for the simple exclusion process work poorly if the initial configuration contains nearby vertices. To avoid this problem, we relate the first collision time in $\{S_t\}_{t\in\mathbb{N}}$ to the first collision time in three other processes which avoid this technical difficulty.

Let $S_1^{(1)} = \mathbf{1}_{S\setminus\{x_2\}}$, let $S_1^{(2)} = \mathbf{1}_{S\setminus\{x_1\}}$ and let $S_1^{(3)} = \mathbf{1}_{\{x_1,x_2\}}$. We let $\{S_t^{(1)}\}_{t\in\mathbb{N}}$, $\{S_t^{(2)}\}_{t\in\mathbb{N}}$ and $\{S_t^{(3)}\}_{t\in\mathbb{N}}$ be simple exclusion processes with these three starting points, coupled to $\{S_t\}_{t\in\mathbb{N}}$ by choosing the same update sequence in Definition 4.3. Let $\tau_{\text{coll}}^{(1)}$, $\tau_{\text{coll}}^{(2)}$ and $\tau_{\text{coll}}^{(3)}$ be their associated collision times, given by the formula

$$\tau_{\text{coll}}^{(\ell)} = \min\{t > 0 : \exists u, v \in \Lambda(L,2) \text{ s.t. } S_t^{(\ell)}[u] = S_t^{(\ell)}[v] = 1 \text{ and } |u-v| = 1\}$$

for $\ell \in \{1,2,3\}$. We note that, under this coupling of the four simple exclusion processes, any single particle in S_t appears in at least one of $S_t^{(1)}$, $S_t^{(2)}$, $S_t^{(3)}$:

$$\{u : S_t[u] = 1\} = \bigcup_{\ell=1}^{3} \{u : S_t^{(\ell)}[u] = 1\}.$$

Furthermore, any *pair* of particles in S_t appears in at least one of $S_t^{(1)}$, $S_t^{(2)}$, $S_t^{(3)}$:

$$\{(u,v) : S_t[u] = S_t[v] = 1\} = \bigcup_{\ell=1}^{3} \{(u,v) : S_t^{(\ell)}[u] = S_t^{(\ell)}[v] = 1\}.$$

Since τ_{coll} and $\{\tau_{\text{coll}}^{(\ell)}\}_{\ell=1}^{3}$ are determined by the positions of pairs of particles, this implies

$$\tau_{\text{coll}} = \min\bigl(\tau_{\text{coll}}^{(1)}, \tau_{\text{coll}}^{(2)}, \tau_{\text{coll}}^{(3)}\bigr). \tag{5.5}$$

Thus, to bound τ_{coll}, it is enough to bound these other collision times. By Theorem 4 of [34], there exists some $\delta_1 = \delta_1(m)$ so that

$$\begin{aligned} \mathbb{P}\left[\tau_{\text{coll}}^{(1)} < \delta_1 \frac{n^2}{\log(n)}\right] &< \frac{1}{\log(n)^2}, \\ \mathbb{P}\left[\tau_{\text{coll}}^{(2)} < \delta_1 \frac{n^2}{\log(n)}\right] &< \frac{1}{\log(n)^2} \end{aligned} \tag{5.6}$$

uniformly in $1 \le i \le m$. By Theorem 4.1 of [18], there exists some $0 < \delta_2, C < \infty$ so that

$$\mathbb{P}\left[\tau_{\text{coll}}^{(3)} < \delta_2 \frac{n^2}{\log(n)}\right] < 1 - \frac{C}{\log(n)}. \tag{5.7}$$

Combining Inequalities (5.6) and (5.7), there exists some $\delta = \delta(m)$ and constant $0 < C < \infty$ so that

$$\mathbb{P}\left[\min\bigl(\tau_{\text{coll}}^{1}, \tau_{\text{coll}}^{2}, \tau_{\text{coll}}'\bigr) < \frac{\delta n^2}{\log(n)}\right] < 1 - \frac{C}{\log(n)}.$$

Combining this with Inequality (5.5) completes the proof. □

We now apply this bound to the KCIP.

For fixed i, let $\{Z_t\}_{t\in\mathbb{N}}$ be a Markov chain with kernel P_i, let $\tau^{(i)} = \tau^{(i)}(1) = \min\{t \in \mathbb{N} : Z_t \in \Omega_i\}$ and let $\tau^{(i+1)} = \tau^{(i+1)}(1) = \min\{t \in \mathbb{N} : Z_t \in \Omega_{i+1}\}$. For $j \in \mathbb{N}$, we define inductively

$$\begin{aligned} \tau^{(i)}(j+1) &= \min\bigl\{t > \tau^{(i+1)}(j) : Z_t \in \Omega_i\bigr\}, \\ \tau^{(i+1)}(j+1) &= \min\bigl\{t > \tau^{(i)}(j) : Z_t \in \Omega_{i+1}\bigr\}. \end{aligned}$$

We then have the corollary:

Corollary 5.4 (Collision from well-spaced configurations). *Fix $m \in \mathbb{N}$ and $1 \le i \le m$. Let $\{Z_t\}_{t\in\mathbb{N}}$ be as above, with initial configuration $Z_1 \in \mathcal{G}(i)$. Then there exists some $\delta = \delta(c, m) > 0$ so that*

$$\mathbb{P}\left[\tau^{(i)} > \delta \frac{n^3}{\log(n)}\right] > \frac{\delta}{\log(n)}.$$

Proof. We consider a simple exclusion process $\{S_t\}_{t\in\mathbb{N}}$ started at $S_1 = Z_1$. We let τ_{coll} be as in Equation (5.4). By analyzing the maximal coupling of S_t and Z_t, it is straightforward to check that there exists some $0 < \gamma, \varepsilon_0 < 1$ so that

$$\mathbb{P}\bigl[\tau^{(i)} > \gamma\varepsilon n^3\bigr] \ge \gamma\mathbb{P}\bigl[\tau_{\text{coll}} > \varepsilon n^2\bigr] \tag{5.8}$$

for all $0 < \varepsilon < \varepsilon_0$. Applying Lemma 5.3 completes the proof. Note that a detailed proof of Inequality (5.8) is given in the first half of the proof of Lemma 7.4 of [32]. □

We are now ready to prove the first of our main bounds on transition times between Ω_i and Ω_{i+1}:

Lemma 5.5. *Fix $0 < r < \infty$. There exists some $C_1 = C_1(c, r)$ so that*

$$\max_{z\in\Omega_i\cup\Omega_{i+1}} \mathbb{E}\bigl[\tau^{(i)}\bigr] \le C_1 n^3 \log(n) \tag{5.9}$$

uniformly in $1 \le i \le r\log(n)$. *There exists some* $C_2 = C_2(c,r)$ *so that*

$$\max_{z\in\Omega_i\cup\Omega_{i+1}} \mathbb{E}\left[\tau^{(i+1)}\right] \le C_2 n^3 \tag{5.10}$$

uniformly in $1 \le i \le r\log(n)$. *There exists some* $C_3 = C_3(c,r)$ *so that*

$$\min_{z\in\Omega_i} \mathbb{P}\left[\tau^{(i+1)} > \frac{n^3}{C_3\log(n)^2}\right] \ge C_3^{-1}. \tag{5.11}$$

uniformly in $1 \le i \le r\log(n)$.

Proof. We begin by proving Inequality (5.10). We will bound from below the probability that the original KCIP $\{X_t\}_{t\in\mathbb{N}}$ travels from $\Omega_{n,i}$ to $\Omega_{n,i+1}$ via a specific sequence of events (see Equation (5.12)) that turn out to describe a "typical" transition between these sets.

To simplify notation, we observe that we can simulate a step of the Markov chain $\{Z_t\}_{t\in\mathbb{N}}$ in terms of the KCIP $\{X_t\}_{t\in\mathbb{N}}$ with starting point $X_1 = Z_1$ according to the following rather inefficient rejection-sampling algorithm:

Definition 5.6 (Coupling of Trace and KCIP). With notation as above, the following is a valid coupling of the KCIP and one step of its trace:

(1) Simulate $\{X_t\}_{t\in\mathbb{N}}$.
(2) Define $\eta = \min\{s > 1 : X_s \in \bigcup_j \Omega_j\}$.
(3) If $X_\eta \in \Omega_i \cup \Omega_{i+1}$, set $Z_2 = X_\eta$. Otherwise, go back to step 1.

We now analyze this algorithm. Recall G_t and $\mathrm{ConnComp}(G_t)$ introduced in the beginning of Section 3. Fix $X_1 \in \Omega_i$ and define the events and random times

$$\begin{aligned}
\mathcal{A} &= \left\{\exists v \in \Lambda(L,2) : X_2 = X_1 \cup \{v\} \text{ and } \left|\left\{u : X_1[u] = 1, |u-v| \le 1\right\}\right| = 1\right\},\\
\kappa &= \inf\{s > 2 : X_s \ne X_{s-1}\},\\
\mathcal{B} &= \left\{|X_\kappa| > |X_2|, \mathrm{ConnComp}(G_\kappa) = \mathrm{ConnComp}(G_1)\right\} \cap \mathcal{A},\\
\zeta &= \inf\{s > \kappa : X_s \ne X_{s-1}\},\\
\mathcal{C} &= \{X_\zeta \in \Omega_{i+1}\} \cap \mathcal{A} \cap \mathcal{B}.
\end{aligned} \tag{5.12}$$

By direct computation,

$$\begin{aligned}
&\mathbb{P}[\mathcal{A}] \ge \frac{c}{n^2},\\
&\mathbb{E}[\mathbf{1}_{\mathcal{B}}|\mathcal{A}] \ge \mathbf{1}_{\mathcal{A}}\left(\frac{c}{2n} - O\left(\frac{r\log(n)}{n^2}\right)\right),\\
&\mathbb{E}[\mathbf{1}_{\mathcal{C}}|\mathcal{A},\mathcal{B}] \ge \mathbf{1}_{\mathcal{A}\cap\mathcal{B}}\left(\frac{1}{4} - O\left(\frac{r\log(n)}{n}\right)\right).
\end{aligned}$$

Combining these bounds, we have

$$\mathbb{P}[Z_1 \in \Omega_{i+1}] \ge \mathbb{P}[\mathcal{C}] \ge \frac{C}{n^3}$$

for some $C = C(r,c) > 0$. This completes the proof of Inequality (5.10).

Inequality (5.9) is proved exactly as the first inequality in Lemma 7.6 of [32], with one small change: the single reference to Theorem 5 of [34] should be replaced by a reference to Theorem 1.1 of [26]. Inequality (5.11) is exactly Lemma 4.1 of [32]. □

Our second main bound on the transition probabilities is:

Lemma 5.7. *For fixed $m \in \mathbb{N}$, there exist constants $C_1 = C_1(m,c)$, $C_2 = C_2(m,c)$, $C_3 = C_3(m,c)$ so that*

$$\min_{z \in \Omega_{i+1}} \mathbb{P}\left[\tau^{(i)}\big(C_2 \log(n)^2\big) > \frac{n^3}{C_1 \log(n)^3}\right] \geq \frac{C_3}{\log(n)}$$

uniformly in $1 \leq i \leq m$.

Proof. Let $\{Z_t\}_{t \in \mathbb{N}}$ be a Markov chain evolving according to P_i, with $Z_1 \sim \mathrm{unif}(\Omega_i)$. Define the measure μ_i on Ω_{i+1} by

$$\mu_i(A) = \mathbb{P}[Z_2 \in A | Z_1 \in \Omega_{i+1}].$$

Recall the definition of $\mathcal{G}(i)$ in Definition 5.3. We note that, by the usual 'coupon collector' problem[6] and the observation that $1 - o(1)$ of the transitions in P_i correspond to adding a single particle in the underlying KCIP (see Definition 5.6 for a precise coupling of P_i and the KCIP which makes this fact clear), we have

$$\mu_i\big(\mathcal{G}(i)\big) = 1 - o(1).$$

Next, let $\{Y_t\}_{t \in \mathbb{N}}$ be a Markov chain evolving according to P_i with any initial state $Y_1 = z \in \Omega_i$ in Ω_i. Combining the hitting and occupation bounds of Lemma 5.5 with the bound on the mixing time $\tau_{\mathrm{mix}}^{(i)}$ given in Lemma 4.2, for all $\varepsilon > 0$ there exists a constant $L = L(\varepsilon, m, c)$ so that

$$\mathbb{P}\big[Y_{\tau^{(i+1)}(L\log(n)^2)} \in \mathcal{G}(i)\big] \geq \mu_i\big(\mathcal{G}(i)\big) - \varepsilon = 1 - \varepsilon - o(1). \tag{5.13}$$

By Corollary 5.4, there exists some $\delta = \delta(m,c)$ so that

$$\mathbb{P}\left[\tau^{(i)}\big(L\log(n)^2 + 1\big) - \tau^{(i)}\big(L\log(n)^2\big) > \delta \frac{n^3}{\log(n)^2} \,\middle|\, Y_{\tau^{(i+1)}(L\log(n)^2)} \in \mathcal{G}(i)\right] \geq \frac{\delta}{\log(n)}.$$

Combining this with Inequality (5.13) completes the proof. □

We are ready to prove the main bound in this section:

Lemma 5.8. *Fix $0 < r < \infty$. Then there exists $C = C(r,c)$ so that*

$$1 - \beta_1(P_i) \geq \frac{C}{n^3 \log(n)^9}$$

uniformly in $1 \leq i \leq r\log(n)$.

Proof. The idea behind this proof is to bound the restriction P_i of the KCIP process to $\Omega_i \cup \Omega_{i+1}$. Using Lemma 2.1 of [31], it is possible to get such a bound by "piecing together" individual bounds on the spectral gap of the trace of the KCIP on Ω_i and Ω_{i+1} by using bounds on the transition times between Ω_i and Ω_{i+1}.

For $T \in \mathbb{N}$, let $N_i(T) = |\{0 \leq t \leq T : Z_t \in \Omega_i\}|$ and $N_{i+1}(T) = |\{0 \leq t \leq T : Z_t \in \Omega_{i+1}\}|$. By Lemmas 5.5 and 5.7, for all $M \in \mathbb{N}$ there exists some $C = C(r,c,M)$ so that

$$\mathbb{P}\big[N_i(T) < Mn^3 \log(n)^3\big] \leq \frac{1}{100} \tag{5.14}$$

for all $T > Cn^3\log(n)^5$, uniformly in $1 \leq i \leq r\log(n)$. The same lemmas imply that for all $m, M \in \mathbb{N}$, there exists some $C = C(m,c,M)$ so that

$$\mathbb{P}\big[N_{i+1}(T) < Mn^3 \log(n)^3\big] \leq \frac{1}{100} \tag{5.15}$$

for all $T > Cn^3\log(n)^9$, uniformly in $1 \leq i \leq m$.

[6] See *e.g.* the survey [12] for exact calculations related to the coupon collector problem.

Let $m_{\max} = 100\max(1, c)$. We note that, for $i \geq m_{\max}$ and all $n > N_0$ sufficiently large, $\frac{\pi(\Omega_i)}{\pi(\Omega_i \cup \Omega_{i+1})} > 0.51$.

Combining the occupation bound in Inequalities (5.14) and (5.15) with the mixing bound in Lemma 4.2, for all $0 < C < \infty$ there exists $M = M(C, c)$

$$\mathbb{P}\big[N_i(T) > C\tau_{n,i}\big] > 0.99, \quad 100m_{\max} < i \leq 2\log(n),$$
$$\mathbb{P}\big[\big\{N_i(T) > C\tau_{n,i}\big\} \cap \big\{N_{i+1}(T) > C\tau_{n,i+1}\big\}\big] > 0.99, \quad 1 \leq i \leq 100m_{\max}$$

for all $T > Mn^3\log(n)^9$. The result now immediately follows from Lemma 2.1 of [31] and the observation that

$$\frac{\pi(\Omega_i)}{\pi(\Omega_i \cup \Omega_{i+1})} > 0.51$$

for all $i > 100m_{\max}$. □

5.3. *Proof of Lemma 5.1*

Applying Theorem 3, with bounds on the individual spectral gaps given by Lemmas 5.2 and 5.8, there exists some $0 < C = C(r, c) < \infty$ so that

$$1 - \beta_1(P_{n,r\log(n)}) \geq \frac{C}{n^3\log(n)^{11}}.$$

Applying the standard bound on the mixing time of a finite Markov chain in terms of its spectral gap (see *e.g.* Theorem 12.3 of [21]) completes the proof of Lemma 5.1.

6. Proof of Theorem 1

Theorem 3 of [32] yields that, there exists some constant $C = C(c)$ so that

$$\tau_{\text{mix}} \geq Cn^3.$$

We now prove the upper bound on τ_{mix}. For fixed $0 < r < \infty$ and $T \in \mathbb{N}$, define the occupation time

$$N(r, T) = \left|\left\{1 \leq t \leq T : X_t \in \bigcup_{i=1}^{\lfloor r\log(n)\rfloor} \Omega_i\right\}\right|.$$

We claim:

Proposition 6.1. *With notation as above, there exists some $r = r_{\max}(c) < \infty$ and $C_1 = C_1(c, r)$, $C_2 = C_2(c, r)$ so that*

$$\mathbb{P}\big[N(r, T) \leq C_1n^3\log(n)^{13}\big] \leq \frac{1}{100} \tag{6.1}$$

for all $T > C_2n^3\log(n)^{14}$.

Proof. Fix ε_0 as in the statement of Theorem 3.1, let $\varepsilon = \frac{1}{2}\varepsilon_0$, and let α, C_G and $\{V_t\}_{t\in\mathbb{N}}$ be as in the statement of Theorem 3.1. Fix $r = \frac{2}{\alpha}C_G$ and define $\mathcal{K} = \{x \in \{0, 1\}^{\Lambda(L,2)} : \sum_{v\in\Lambda(L,2)} x[v] \leq r\log(n)\}$. Let $\tau_{\text{start}} = \inf\{t \in \mathbb{N} : X_t \in \mathcal{K}\}$ and fix $k \in \mathbb{N}$. By Theorem 3.1,

$$\mathbb{E}[V_{k\varepsilon n^3\log(n)}\mathbf{1}_{\tau_{\text{start}} > k\varepsilon n^3\log(n)}] \leq \left(1 - \frac{1}{2}\alpha\right)^k V_1,$$

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

$$\begin{aligned}\mathbb{P}\big[\tau_{\text{start}} > k\varepsilon n^3 \log(n)\big] &\le \mathbb{P}[V_{k\varepsilon n^3}\mathbf{1}_{\tau_{\text{start}} > k\varepsilon n^3} > 1] \\ &\le n\left(1 - \frac{1}{2}\alpha\right)^k. \end{aligned} \tag{6.2}$$

Fix $T \in \mathbb{N}$ and constants $0 < C_1, C_2 < \infty$. 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 (6.2), the Markov property and Lemma 7.1 of [32],

$$\begin{aligned}&\mathbb{P}\left[\sum_{t=1}^{C_1 n^3 \log(n)^{14}} \mathbf{1}_{X_t\in\mathcal{K}} > C_2 n^3 \log(n)^{13}\right] \\ &\ge \mathbb{P}\left[\sum_{t=1}^{C_1 n^3 \log(n)^{14}} \mathbf{1}_{X_t\in\mathcal{K}} > C_2 n^3 \log(n)^{13} | \tau_{\text{start}} < T\right]\mathbb{P}[\tau_{\text{start}} < T] \\ &= \mathbb{P}[\tau_{\text{start}} < T]\sum_{s=1}^{T}\mathbb{P}\left[\sum_{t=s}^{C_1 n^3 \log(n)^{14}} \mathbf{1}_{X_t\in\mathcal{K}} > C_2 n^3 \log(n)^{13} | \tau_{\text{start}} = s\right]\mathbb{P}[\tau_{\text{start}} = s | \tau_{\text{start}} \le T] \\ &\ge \left(1 - n\left(1 - \frac{1}{2}\alpha\right)^{\lfloor \frac{T}{\varepsilon n^3 \log(n)}\rfloor}\right)\mathbb{P}\left[\sum_{t=T}^{C_1 n^3 \log(n)^{14}} \mathbf{1}_{X_t\in\mathcal{K}} > C_2 n^3 \log(n)^{13} | \tau_{\text{start}} \le T\right] \\ &\ge \left(1 - n\left(1 - \frac{1}{2}\alpha\right)^{\lfloor \frac{T}{\varepsilon n^3 \log(n)}\rfloor}\right)\left(1 - \lceil \varepsilon n^3 \log(n)\rceil \mathbb{P}\left[\sum_{j=1}^{C_2 \log(n)^{13}} Z_j \le C_1 \log(n)^{14} - \frac{T}{\varepsilon n^3 \log(n)}\right]\right). \end{aligned} \tag{6.3}$$

Choosing $T = \lfloor \frac{C_1}{2} n^3 \log(n)^{14} \rfloor$, we have for C_1 sufficiently large that

$$\mathbb{P}\left[\sum_{j=1}^{C_2 \log(n)^{13}} Z_j \le C_1 \log(n)^{14} - \frac{T}{\varepsilon n^3 \log(n)}\right] = 1 - o\big(n^{-10}\big)$$

by a standard concentration inequality for geometric random variables. Combining this with the calculation (6.3) completes the proof. □

The upper bound on τ_{mix} now follows immediately from Lemma 2.1 of [31] as explained in *Step* 1 of Section 2, with the bound on the occupation time given by Inequality (6.1) and the bound on the maximal mixing time given by Theorem 5.1.

References

[1] D. Aldous. Open problems, 2017. Available at https://www.stat.berkeley.edu/~aldous/Research/OP/index.html; last accessed 7-June-2017.

[2] M. Ben and Y. Peres. Evolving sets, mixing and heat kernel bounds. *Probab. Theory Related Fields* **133** (2) (2005) 245–266. MR2198701

[3] O. Blondel, N. Cancrini, F. Martinelli, C. Roberto and C. Toninelli. Fredrickson–Andersen one spin facilitated model out of equilibrium. *Markov Process. Related Fields* **19** (2013) 383–406. MR3156958

[4] N. Cancrini, F. Martinelli, C. Roberto and C. Toninelli. Kinetically constrained models. In *New Trends in Mathematical Physics* 741–752, 2009.

[5] P. Chleboun, A. Faggionato and F. Martinelli. Time scale separation and dynamic heterogeneity in the low temperature East model. *Comm. Math. Phys.* **328** (3) (2014) 955–993. MR3201217

[6] P. Chleboun, A. Faggionato and F. Martinelli. Mixing time and local exponential ergodicity of the East-like process in $\mathbb{Z}^d$. *Ann. Fac. Sci. Toulouse Math. (6)* **24** (4) (2015) 717–743. MR3434253

[7] P. Chleboun, A. Faggionato and F. Martinelli. Relaxation to equilibrium of generalized East processes on $\mathbb{Z}^d$: Renormalization group analysis and energy-entropy competition. *Ann. Probab.* **44** (3) (2016) 1817–1863. MR3502595

[8] P. Chleboun and F. Martinelli. Mixing time bounds for oriented kinetically constrained spin models. *Electron. J. Probab.* **18** (2013) 1–9. MR3084571
[9] P. Chleboun and A. Smith. Mixing of the square plaquette model on a critical length scale. Preprint, 2018. Available at arXiv:1807.00634.
[10] P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika* **60** (3) (1973) 581–588. MR0343950
[11] M. Dyer, L. Goldberg, M. Jerrum and R. Martin. Markov chain comparison. *Probab. Surv.* **3** (2006) 89–111. MR2216963
[12] P. Flajolet, D. Gardy and L. Thimonier. Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Appl. Math.* **39** (3) (1992) 207–229. MR1189469
[13] G. Fredrickson and H. Andersen. Kinetic Ising model of the glass transition. *Phys. Rev. Lett.* **53** (1984) 1244–1247.
[14] G. Fredrickson and H. Andersen. Facilitated kinetic Ising models and the glass transition. *J. Chem. Phys.* **83** (1985) 5822–5831.
[15] J. Garrahan, P. Sollich and C. Toninelli. *Dynamical Heterogeneities in Glasses, Colloids and Granular Media*, Chapter 10. Oxford University Press, London, 2011.
[16] N. Hamamuki. A discrete isoperimetric inequality on lattices. *Discrete Comput. Geom.* **52** (2) (2014) 221–239. MR3249380
[17] R. Holley and T. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3** (4) (1975) 643–663. MR0402985
[18] N. Jain and S. Orey. On the range of random walk. *Israel J. Math.* **6** (4) (1968) 373–380. MR0243623
[19] M. Jerrum, J.-B. Son, P. Tetali and E. Vigoda. Elementary bounds on Poincare and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Probab.* **14** (4) (2004) 1741–1765. MR2099650
[20] G. Kordzakhia and S. Lalley. Ergodicity and mixing properties of the Northeast model. *J. Appl. Probab.* **43** (3) (2006) 782–792. MR2274800
[21] D. Levin, Y. Peres and E. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2009. MR2466937
[22] N. Madras and D. Randall. Markov chain decomposition for convergence rate analysis. *Ann. Appl. Probab.* **12** (2002) 581–606. MR1910641
[23] F. Martinelli and C. Toninelli. Kinetically constrained spin models on trees. *Ann. Appl. Probab.* **23** (5) (2013) 1967–1987. MR3134727
[24] F. Martinelli and C. Toninelli. Towards a universality picture for the relaxation to equilibrium of kinetically constrained models. *Ann. Probab.* **47** (1) (2019) 324–361. MR3909971
[25] R. Oliveira. Mixing and hitting times for finite Markov chains. *Electron. J. Probab.* **17** (70) (2012) 1–12. MR2968677
[26] R. Oliveira. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.* **364** (4) (2012) 2109–2128. MR2869200
[27] J. Palacios and P. Tetali. A note on expected hitting times for birth and death chains. *Statist. Probab. Lett.* **30** (1996) 119–125. MR1416997
[28] Y. Peres and P. Sousi. Mixing times are hitting times of large sets. *J. Theoret. Probab.* **9** (2013) 459–510. MR3370663
[29] D. Persi and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** (3) (1993) 696–730. MR1233621
[30] D. Persi and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** (1996) 695–750. MR1410112
[31] N. Pillai and A. Smith. Elementary bounds on mixing times for decomposable Markov chains. *Stochastic Process. Appl.* **127** (9) (2017) 3068–3109. MR3682124
[32] N. S. Pillai and A. Smith. Mixing times for a constrained Ising process on the torus at low density. *Ann. Probab.* **45** (2) (2017) 1003–1070. MR3630292
[33] A. Smith. Comparison theory for Markov chains on different state spaces and application to random walk on derangements. *J. Theoret. Probab.* **28** (4) (2015) 1406–1430. MR3422936
[34] J. Theodore Cox. Coalescing random walks and voter model consensus times on the torus in $\mathbb{Z}^d$. *Ann. Probab.* **17** (4) (1989) 1333–1366. MR1048930