

PARTICLE DENSITY IN DIFFUSION-LIMITED ANNIHILATING SYSTEMS

BY TOBIAS JOHNSON^{1,a}, MATTHEW JUNGE^{2,b},
HANBAEK LYU^{3,c}, AND DAVID SIVAKOFF^{4,d}

¹Mathematics, CUNY College of Staten Island, ^aTobias.Johnson@csi.cuny.edu

²Mathematics, CUNY Baruch College, ^bMatthew.Junge@baruch.cuny.edu

³Mathematics, University of Wisconsin - Madison, ^chlyu@math.wisc.edu

⁴Mathematics and Statistics, Ohio State University, ^ddsivakoff@stat.osu.edu

Place an A -particle at each site of a graph independently with probability p and otherwise place a B -particle. A - and B -particles perform independent continuous time random walks at rates λ_A and λ_B , respectively, and annihilate upon colliding with a particle of opposite type. Bramson and Lebowitz studied the setting $\lambda_A = \lambda_B$ in the early 1990s. Despite recent progress, many basic questions remain unanswered when $\lambda_A \neq \lambda_B$. For the critical case $p = 1/2$ on low-dimensional integer lattices, we give a lower bound on the expected number of particles at the origin that matches physicists' predictions. For the process with $\lambda_B = 0$ on the integers and on the bidirected regular tree, we give sharp upper and lower bounds for the expected total occupation time of the root at and approaching criticality.

1. Introduction. We consider two-type diffusion limited annihilating systems (DLAS) on integer lattices and directed regular trees. Initially every vertex has a particle that is independently of type A with probability p and otherwise is of type B . In continuous time, A -particles perform simple random walk at rate λ_A and B -particles at rate λ_B . If two particles of opposite types collide, both are removed from the system. Particles annihilate pairwise and there is no limit to the number of like particles that can occupy a single site. See Figure 1 for a depiction of a simulation.

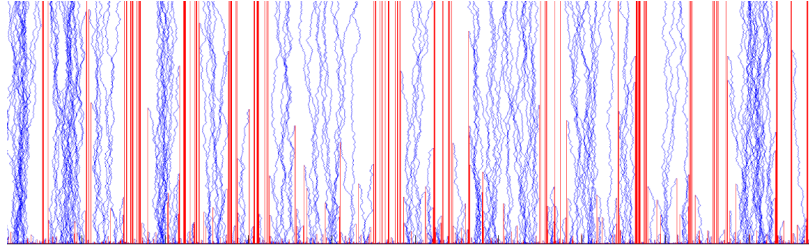


FIG 1. Space-time representation of a DLAS with $p = 1/2$ and $\lambda_B = 0$ on the path with 2000 vertices. Vertical lines are B -particles. Time runs from bottom to top.

Physicists have been interested in this process as a model for irreversible reactions with mobile particles since the papers [34] and [42]. The mean-field prediction for the model is that the density of both particle types should decay at rate t^{-1} if $p = 1/2$, while the density of the less common particle type should decay exponentially if $p \neq 1/2$. It was widely observed in

MSC2020 subject classifications: Primary 60K35.

Keywords and phrases: Interacting Particle System, Critical Behavior.

the physics literature that while these predictions were correct in high dimension, the model followed different asymptotics in low dimension. But as Bramson and Lebowitz noted in [9], “the answers given in that literature do not always agree.” They then rigorously determined the asymptotics of the model on $\mathbb{Z}^d$ for all values of d when $\lambda_A = \lambda_B$ in [10]. Note that they consider slightly different initial conditions than our independent Bernoulli field. In their model, the number of A - and B -particles at each site are given by independent Poisson fields with intensities r_A and r_B for A - and B -particles, respectively. Let ρ_t denote the expected number of A -particles at the origin at time t . Bramson and Lebowitz showed that when $r_A = r_B$ (analogous to $p = 1/2$ in our model),

$$(1) \quad \rho_t \asymp \begin{cases} t^{-d/4}, & d \leq 3 \\ t^{-1}, & d \geq 4 \end{cases}$$

using the notation $f(t) \asymp g(t)$ to denote that $f(t)/g(t)$ is bounded from above and below by positive constants (which may depend on d and p) for all t . On the other hand, when $p < 1/2$,

$$(2) \quad e^{-c_d \psi_d(p,t)} \leq \rho_t \leq e^{-C_d \psi_d(p,t)},$$

for all large t where c_d and C_d are positive constants depending only on d and

$$\psi_d(p,t) = \begin{cases} \frac{(1-2p)^2}{1-p} \sqrt{t}, & d = 1 \\ (1-2p) \frac{t}{\log t}, & d = 2 \\ (1-2p)t, & d \geq 3. \end{cases}$$

This confirms that the model deviates from the mean-field behavior in dimensions $d \leq 3$ when the initial particle densities are equal and in dimensions $d \leq 2$ when the densities are unequal.

The closely related system *annihilating random walk* was already known not to exhibit mean-field behavior in low dimension. This process typically starts with one particle per site with particles performing independent random walks at rate 1. Any collision results in mutual annihilation. Arratia proved in [2] that β_t , the density of particles at the origin at time t , satisfies

$$(3) \quad \beta_t \asymp \begin{cases} t^{-1/2}, & d = 1 \\ t^{-1} \log t, & d = 2 \\ t^{-1}, & d \geq 3 \end{cases}.$$

Arratia’s work built on similar results from Bramson and Griffeath [8] for *coalescing random walk*, in which particles coalesce upon colliding (equivalently one particle is annihilated in each collision). Coalescing random walk has nice monotonicity properties and also enjoys a dual process known as the *voter model*. Bramson and Griffeath [8] used this dual process to obtain sharp asymptotic for particle density in coalescing random walk, and Arratia [2] used a coupling between coalescing and annihilating random walk to show that the particle density in the annihilating system is asymptotically the half of the coalescing system. This leads to the exponents in (3). There is no known tractable dual process for DLAS or coupling to well-known processes.

The asymptotics given in (1) and (2) were also conjectured to hold for different jump rates [30, 28, 27]. However, a lack of symmetry makes these dynamics more difficult to analyze. For example, Bramson and Lebowitz analyzed DLAS with a coupling that first ignores the type of a particle and later reveals it. Such a coupling is valid only when $\lambda_A = \lambda_B$. Another key result, which appears to hold only in the same-speed setting, is [10, Lemma 2.1]. This states that the average value of any convex function of the difference of the numbers of A - and

B -particles in a region at time t will be at least as large for DLAS compared to the process with no annihilation. A bound on ρ_t when $\lambda_A \neq \lambda_B$ was proven in [13] by Cabezas, Rolla, and Sidoravicius: for any choice of jump rates $\lambda_B \geq 0$ and $\lambda_A > 0$, it holds that $\rho_t \geq C/t$ on a large class of transitive, unimodular graphs (including $\mathbb{Z}^d$) when the particle densities are in balance. In particular, this is in line with (1) for $d \geq 4$. The same authors in [12] proved that when $\lambda_B = 0$ and A -particles move with drift, the number of visits to the origin in time t is of order $\sqrt{t}$. It turns out, however, that the asymptotic behavior for $p < 1/2$ is different when $\lambda_B = 0$ from when $\lambda_A = \lambda_B > 0$. Damron, Lyu and Sivakoff show that when $\lambda_B = 0$ and A -particles move as discrete-time symmetric random walks, then $\rho_t = \exp(-O(t^{d/(d+2)}))$, which is distinct from (2) for all dimensions $d \geq 1$ [18]. This cast some doubt on whether the same asymptotics as in (1) and (2) should always be expected when $\lambda_B < \lambda_A$.

Some other related results include: Cristali, Jiang, Junge, Kassem, Sivakoff, and York considered a discretized version of DLAS on finite graphs and studied the time to extinguish all particles [17]. Ahlberg, Griffiths, and Janson studied the critical behavior of an two-type annihilating system of branching random walks [1]. Dauvergne and Sly recently studied a variant of DLAS in which A - and B -particles move at different rates but, rather than collisions resulting in mutual annihilation, B -particles are converted to A -particles upon contact [20]. Bahl, Barnet, Junge, and Johnson proved in [6] that the occupation time of a subset of vertices by A -particles in DLAS is monotonic as the initial configuration is augmented in the increasing convex order, a nonstandard stochastic order that rewards volatility.

Cabezas, Rolla, and Sidoravicius further proved that DLAS undergoes a phase transition from infinite visits by A -particles to the root (recurrence) when $p \geq 1/2$ to only finitely many visits (transience) for $p < 1/2$ [13]. This built on previous recurrence/transience results by the same authors [12] for the special case $\lambda_B = 0$, which they named the *particle-hole model*. An Abelian property ensures their results also hold in discrete time. Damron, Gravner, Junge, Lyu, and Sivakoff produced similar results for the case $\lambda_B = 0$ with A -particles performing discrete time random walk and also derived some quantitative estimates on the expected number of particles to visit the origin in [19]. Inspired by recent results in parking on random graphs [29, 23], the authors named this setting the *parking process*.

As we were writing up our results, Przykucki, Roberts, and Scott released a paper concerning the parking process on the integers [36]. For the case $p = 1/2$, they proved lower and upper bounds on the expected occupation time of the root, $\mathbb{E}_p V_t = \int_0^t \rho_s ds$, that matched the conjectured behavior up to a sublogarithmic factor:

$$(4) \quad t^{3/4}(\log t)^{-1/4} \lesssim \mathbb{E}_{1/2} V_t \lesssim t^{3/4}.$$

Here, the occupation time V_t is the integral of the number of A -particles at the origin at time s from $s = 0$ to t , so it accounts for multiple visits of the same particle to the origin up to time t . The notation $f(t) \lesssim g(t)$ means $f(t) = O(g(t))$. Addressing a conjecture from [19] concerning the rate of growth of $\mathbb{E}_p V_\infty$ as p approaches $1/2$, they further proved that as $p \uparrow 1/2$

$$(5) \quad \mathbb{E}_p V_\infty \lesssim (1 - 2p)^{-6}.$$

Using different techniques, we prove a stronger version of this estimate. We also provide several other estimates on ρ_t , which we summarize below and then state more precisely.

For $d \leq 3$, we give a lower bound of $\rho_t \geq (t \log t)^{-d/4}$ for the $p = 1/2$ case (see Theorem 1.1). This is consistent with the behavior with equal jump rates in (1) and with $\lambda_B = 0$ in (4) for the case $\lambda_B = 0$ in discrete time on $\mathbb{Z}$. Our work and [36] are the first confirmations of deviation from mean-field behavior with nonequal jump rates. For $d = 1$ and $\lambda_B = 0$, we also give an upper bound on the total occupation time for a site that essentially confirms that the asymptotics of (1) hold in this case (Theorem 1.2). This agrees with the upper bound in [36]

except ours is proven in continuous time. Addressing (5), we prove that, up to a logarithmic factor, $\mathbb{E}_p V_\infty$ grows like $(1 - 2p)^{-3}$ (Theorems 1.3 and 1.4).

Our final results investigate the high-dimensional behavior. Bramson and Lebowitz showed that DLAS has the mean-field density decay of t^{-1} on $\mathbb{Z}^d$ for $d \geq 4$ in the case of equal jump rates. We consider the model with $\lambda_B = 0$ on a directed regular tree. We give upper and lower bounds of order $\log t$ on the cumulative occupation time for the $p = 1/2$ case, and we give upper and lower bounds as $p \uparrow 1/2$ that agree up to constants (Theorems 1.5 and 1.6). To the best of our knowledge, this is the only instance in which mean-field behavior has been proven to occur with nonequal jump rates on an infinite graph.

1.1. Statement of results. We consider the two-type DLAS on a given rooted graph where each vertex initially contains exactly one particle, which has type A with probability p , with jump rates for the two particle types given by λ_A and λ_B . Without loss of generality, we can take one of the jump rates to be 1. We take $\lambda_A = 1$ in all results except for Theorem 1.3, where doing so would result in a needless loss of generality.

Let N_t be the number of A -particles at the root at time t . Let $\rho_t = \mathbb{E}N_t$, which we refer to as the density of A -particles at time t . Finally, let $V_t = \int_0^t N_s ds$, the aggregate time spent by A -particles at the root up to time t . All of our results concern the asymptotic behavior of ρ_t and $\mathbb{E}V_t$ on transitive graphs, rendering the choice of root irrelevant.

Our first result is a general lower bound on density with particle types in balance, confirming the lower bounds (up to a logarithmic factor) of the Bramson–Lebowitz asymptotics (1) in low dimension.

THEOREM 1.1. *Let $\lambda_A = 1$ and $0 \leq \lambda_B \leq 1$. On $\mathbb{Z}^d$ with $d \leq 3$ and $p = 1/2$ there exists a constant $C > 0$ that does not depend on λ_B such that*

$$\rho_t \geq C(t \log t)^{-d/4}$$

for all sufficiently large t

Next, we provide an upper bound on $\mathbb{E}V_t$ in dimension 1 with $\lambda_B = 0$. Combined with Theorem 1.1 and the fact that $\mathbb{E}V_t = \int_0^t \rho_s ds$, it strongly supports the conjecture that $\rho_t \asymp t^{-1/4}$.

THEOREM 1.2. *Let $\lambda_A = 1$ and $\lambda_B = 0$. On $\mathbb{Z}$ with $p = 1/2$ there exists $C > 0$ such that*

$$\mathbb{E}V_t \leq Ct^{3/4}$$

for all sufficiently large t .

Next, we consider the critical exponent of $\mathbb{E}_p V_\infty$ as $p \uparrow 1/2$.

THEOREM 1.3. *Let $0 < \lambda_A \leq 1$ and $0 \leq \lambda_B \leq 1$. On $\mathbb{Z}^d$ for $d \leq 3$ there exists a constant $C > 0$ such that*

$$\mathbb{E}_p V_\infty \geq C \frac{(1 - 2p)^{-(4/d)+1}}{-\log(1 - 2p)}$$

for all $1/4 < p < 1/2$.

THEOREM 1.4. *Let $\lambda_A = 1$ and $\lambda_B = 0$. On $\mathbb{Z}$, there exists $C > 0$ such that*

$$\mathbb{E}_p V_\infty \leq C(1 - 2p)^{-3}$$

for all $p < 1/2$.

For $d = 1$ with $\lambda_B = 0$, Theorems 1.3 and 1.4 show that

$$\frac{(1 - 2p)^{-3}}{-\log(1 - 2p)} \lesssim \mathbb{E}_p V_\infty \lesssim (1 - 2p)^{-3},$$

determining the critical exponent up to logarithmic terms. While our upper bound on $\mathbb{E}_p V_\infty$ with $p < 1/2$, $d = 1$, and $\lambda_B = 0$ is sharp (at least up to logarithmic factors), we mention that for $d \geq 2$ and $\lambda_B < \lambda_A$ it remains open to show the much weaker statement that $\mathbb{E}_p V_\infty < \infty$ for all $p < 1/2$.

Our final results concern the mean-field behavior $\rho_t \asymp t^{-1}$. This was first proven by Bramson and Lebowitz on lattices of dimension 4 and higher in the case of equal particle densities and jump rates. While it is believed to hold for all jump rates, it has not been proven to occur with unequal jump rates on any graph. We consider DLAS with $\lambda_B = 0$ on $\vec{\mathcal{T}}_{2d}$, the bidirected $2d$ -regular tree, which is the $2d$ -regular tree where each vertex has d edges directed away from it and d edges directed towards it. We prove that $\mathbb{E} V_t$ diverges logarithmically, which strongly suggests the mean-field density decay. The rationale for working on $\vec{\mathcal{T}}_{2d}$ is that the A -particles approaching the root from different branches of the tree evolve independently, but analysis remains difficult even with this advantage.

THEOREM 1.5. *Let $\lambda_A = 1$ and $\lambda_B = 0$. For some positive absolute constants c and C , it holds for all $d \geq 2$ on $\vec{\mathcal{T}}_{2d}$ with $p = 1/2$ that*

$$c \log t \leq \mathbb{E} V_t \leq C \log t$$

for all large t .

Finally, as in Theorems 1.3 and 1.4, we investigate how quickly $\mathbb{E}_p V_\infty$ diverges as $p \uparrow 1/2$.

THEOREM 1.6. *Let $\lambda_A = 1$ and $\lambda_B = 0$. For some positive absolute constants c , C , and η , it holds on $\vec{\mathcal{T}}_{2d}$ for all $\frac{1}{2} - \eta < p < \frac{1}{2}$ and $d \geq 2$ that*

$$c \log(1 - 2p)^{-1} \leq \mathbb{E}_p V_\infty \leq C \log(1 - 2p)^{-1}.$$

1.2. Definitions and notation. DLAS on a graph with vertex set $\mathcal{V}$ is a continuous-time Markov process $\zeta_t = (\zeta_t(v))_{v \in \mathcal{V}}$ on state space $\mathbb{Z}^{\mathcal{V}}$. The quantity $|\zeta_t(v)|$ denotes the number of particles at site v at time t . The sign of $\zeta_t(v)$ is positive if these particles are of type A and negative if they are of type B . The dynamics of the process are as described earlier: particles of types A and B jump at rates λ_A and λ_B , respectively; a particle at vertex u takes its next step to v with probability $K(u, v)$, where K is a given random walk kernel; and when a particle jumps onto a site with a particle of the opposite type, both particles are instantly annihilated. The infinitesimal generator corresponding to this description is given explicitly in [13, Section 2]. A graphical construction proving that such a Markov process exists is sketched in [10] and given in [13].

If we do not mention the random walk kernel K specifically, then we take it to be the nearest-neighbor simple random walk kernel on the given graph. Our default initial conditions consist of one particle per site, each of which independently is an A -particle with probability p or a B -particle with probability $1 - p$. We will also frequently consider these initial conditions restricted to a subset of the graph, meaning that rest of the graph is initially devoid of particles.

In the previous section, we defined the quantities N_t , ρ_t , and V_t . We will also use the notation $D(H)$ to denote the *discrepancy* between A -particles and B -particles on a subgraph H , defined as the number of A -particles minus the number of B -particles initially placed in

a subgraph H in a given instance of DLAS. We let $\mathbb{T}_r^d$ denote the d -dimensional torus of radius r , which has vertex set $(-r, r]^d \cap \mathbb{Z}^d$ and nearest neighbor edges with the canonical identifications of opposite sides. We denote by $\mathbf{0}$ the origin of the lattices and torus, and also the root of the bidirectional tree.

1.3. Overview of proofs. Our results rely on a variety of couplings that allow us to make comparisons to modified versions of the systems. We sketch the main ideas below.

Theorems 1.1 and 1.3, lower bounds for $d = 1, 2, 3$. The idea behind Theorem 1.1 is to compare ρ_t to the particle density of $\mathbb{T}_{2r}^d$ with $r = C\sqrt{t \log t}$, which we denote as $\bar{\rho}_t$. The width of the torus is chosen so that number of particles at the origin up to time t is unlikely to be affected by the particles beyond distance r in the processes on $\mathbb{Z}^d$ and on $\mathbb{T}_{2r}^d$ (see Lemma 2.3). So it is enough to estimate $\bar{\rho}_t$.

As the number of vertices in $\mathbb{T}_{2r}^d$ is on the order of r^d , standard central limit theorem estimates give that with positive probability, there are $r^{d/2}$ more A - than B -particles in the initial configuration. Since the torus is a finite graph, these surplus A -particles are never annihilated. Translation invariance ensures that

$$\bar{\rho}_t \gtrsim \frac{r^{d/2}}{|\mathbb{T}_{2r}^d|} \gtrsim (t \log t)^{-d/4}.$$

The critical exponent bound given in Theorem 1.3 uses the same idea, but optimizes the size of the torus as a function of p and uses more precise estimates for the A -particle surplus in the initial configuration. Similar ideas are used when making estimates on the *correlation length* in first passage percolation [5]. Note that Bramson and Lebowitz [10] also made use of fluctuations in the initial configuration when studying the symmetric speeds case. See the heuristic at [10, p. 4].

Theorems 1.2 and 1.4, upper bounds for $d = 1$. The starting point is to consider a sequential version of DLAS in which all B -particles are initially present and A -particles are released one at a time according to an arbitrary prescribed ordering. The first A -particle travels until it annihilates with a B -particle or time t elapses. Then, in this new environment, the next A -particle is released and does the same, and so on. For $\lambda_B = 0$, Lemma 2.4 establishes that the occupation time of the root is stochastically larger in this variant than in the original process. Moreover, the subadditivity result in Lemma 2.5 reduces the problem to studying the one-sided version of DLAS with particles only at the positive integers.

For Theorem 1.2, we run the sequential process on the half-line, releasing A -particles in order of their distance from the origin. We show that an A -particle starting at k has probability $O(k^{-1/2})$ of visiting the origin in time t (Lemma 4.2). Summing over all k from 1 to $\sqrt{t}$ and using random walk concentration bounds to bound the contributions of particles starting beyond distance $\sqrt{t}$, we obtain a bound of $O(t^{1/4})$ on the number of distinct particles visiting the origin in the half-line process. If a particle visits the origin, we expect it to spend at most time $O(t^{1/2})$ up to time t there, by basic properties of random walk, yielding the $O(t^{3/4})$ bound for the one-sided sequential process.

The sequential release of particles is essential to the proof of Lemma 4.2, which gives a bound on the probability of A -particles visiting the origin. The main idea is that for an A -particle starting at k to reach $\mathbf{0}$, the A -particles in $[1, k-1]$ must have already annihilated all B -particles that were initially there. The typical surplus of the A -particles against the B -particles in $[1, k-1]$ is $O(k^{1/2})$, and the probability of the A -particle at k reaching $\mathbf{0}$ is maximized if all the surplus A -particles annihilate B -particles to the right of position k . We then estimate the chance of the A -particle at k reaching $\mathbf{0}$ using a refined “gambler’s ruin”

estimate (Proposition A.4), in which we bound the chance of the particle hitting $\mathbf{0}$ by time t and before visiting the first remaining B -particle located approximately $k^{1/2}$ steps to its right.

To prove Theorem 1.4, as in the proof of Theorem 1.2 we use a gambler's ruin approach to bound the time spent at $\mathbf{0}$ by an A -particle starting from position k in terms of the number of surplus A -particles against B -particles on $[0, k]$. When we sum this bound over all k , we obtain a bound on $\mathbb{E}V_t$ in terms of the expected area underneath the positive excursions of a p -biased random walk for $p < 1/2$, which we then compute.

It seems challenging to us to generalize our approach to higher dimensional lattices, because there is no longer a simple way to control the probability an A -particle at distance k reaches the origin. One would have to understand the spatial correlations between unvisited B -particles as the process evolves. These correlations may be significant (see [19, Figure 2].) In the $\lambda_B > 0$ case, it also appears difficult to make analogous estimates even in $d = 1$, since the coupling of the sequential version of DLAS with the usual one depends on B -particles remaining still.

Theorems 1.5 and 1.6, results on bidirected tree $\vec{\mathcal{T}}_{2d}$. These results are proven rather differently from the other estimates. To avoid some technical issues in this explanation, consider DLAS in discrete time. Let W_n denote the number of visits to the root by A -particles in n steps of this process and let $v_1, \dots, v_d$ be the vertices whose out-edges lead to the root of $\vec{\mathcal{T}}_{2d}$. We can express W_{n+1} in terms of the number of visits to $v_1, \dots, v_d$ in time n . From the self-similarity of the tree, each of these quantities is an independent copy of W_n . We thus obtain a distributional equation (28) giving the law of W_{n+1} in terms of the law of W_n . Analysis of this equation shows that the growth of the mean of W_n depends on its concentration (see Lemma 5.5).

We then prove concentration and anticoncentration bounds for W_n . The anticoncentration bound (Lemma 5.21) proves the lower bounds in Theorems 1.5 and 1.6. Note that Cabezas, Rolla, and Sidoravicius show a general result [13, Theorem 4] that $\rho_t = \Omega(1/t)$ when $p = 1/2$ on generously transitive graphs with reflectable jump distributions. The lower bound in Theorem 1.5 of matching order, however, does not follow from this general result, since the jump distribution on directed bidirectional tree is not reflectable. The upper bounds on $\mathbb{E}W_n$ are a consequence of the concentration bound on W_n , which we prove with the technique of size-bias couplings. The idea of this technique is that stochastic inequalities between a random variable and its size-bias transform lead to concentration inequalities for the random variable. The problem then turns to computing the size-bias transform of W_n and showing that it does not differ from W_n by too much. To size-bias a sum of independent terms, one chooses a single term at random to bias, leaving the others unaffected (hence the title of the survey paper [4]). Thus, to size-bias W_n , we bias the number of root visits coming from one of its children. Recursively, this creates a ray on which the process is altered. The result is something like placing an extra A -particle at every vertex along this ray, and concentration then follows from showing that this adds only $O(1)$ visits to the root. The actual details are more complicated; see Lemma 5.17 and Proposition 5.18 and the discussion thereafter.

These size-biasing techniques seem novel to us in the context of particle systems. To put them in context, size-biasing has a long history in Stein's method for distributional approximation (see [38, Sections 3.4 and 4.3]). More recently, size-biasing methods have been developed for proving concentration [22, 3, 16]. On a different track, a technique of creating a spine with altered behavior to bias a statistic of a Galton–Watson tree was used in [32] to prove the Kesten–Stigum Theorem (other good references are [33, Chapter 12] and [4, Section 7]). This technique was later used to prove many results on branching processes; for example, see [24] and [40, Chapters 4 and 5]. These two uses of size-biasing, Stein's method

and spine techniques, finally met in [35] where Stein's method together with the construction from [32] is used to prove a quantitative version of Yaglom's theorem on critical Galton–Watson trees. Of all uses of size-biasing, the most relevant to this paper is used in [25] to analyze the Derrida–Retaux model from statistical physics. This model is essentially DLAS with $\lambda_B = 0$ but on a directed rather than bidirected tree; see [21, 14, 15].

1.4. Organization. Section 2 contains statements of some important lemmas as well as descriptions of variants of DLAS that we relate to the original process. Section 3 contains the proofs of these lemmas. We prove our main results for DLAS on lattices (Theorems 1.1, 1.2, 1.3, and 1.4) in Section 4. In Section 5, we prove our main results on DLAS on bidirected trees (Theorems 1.5 and 1.6). The appendix contains some useful random walk estimates.

2. Key lemmas. In this section, we present a toolkit of lemmas for DLAS, whose proofs are given in Section 3. A reader more interested in how the lemmas are applied may safely read the statements and skip ahead to Section 4. We start with a quick overview of the lemmas and where they are used.

- Lemma 2.1 asserts that removing A -particles results in fewer visits to the root by A -particles. We use this only in one place in Section 5.
- Lemma 2.2 shows that ρ_t does not change much when we remove all particles beyond distance $\sqrt{Ct \log t}$ from the root of the torus or lattice or beyond distance ct from the root of the bidirected tree. We use the lemma in the proofs of Lemma 2.3, Theorem 1.5, and Theorem 1.6.
- In Lemma 2.3 we relate DLAS on $\mathbb{Z}^d$ to DLAS on a torus. We use this to prove the lower bounds on the lattice, Theorems 1.1 and 1.3.
- Lemma 2.4 shows that when $\lambda_B = 0$, the occupation time of $\mathbf{0}$ in DLAS is stochastically larger if we release A -particles one at a time, letting each one run for a fixed time t or until annihilation before running the next. This result is crucial for proving Theorems 1.2 and 1.4, our upper bounds for one-dimensional DLAS.
- Lemma 2.5 is a subadditivity result saying that when $\lambda_B = 0$, dividing the particles of DLAS into two sets and running them as two independent systems makes the combined number of visits to the root stochastically larger. We need this for Theorems 1.2 and 1.4.

Now we state the lemmas precisely. Several statements are for DLAS on transitive unimodular graphs i.e., graphs with a transitive unimodular automorphism group. This includes the graphs $\mathbb{Z}^d$, $\mathbb{T}_r^d$, $\vec{\mathbb{T}}_{2d}$ to which our theorems apply. We work at this level of generality when possible, since [13, Section 6] proved that the graphical construction of DLAS is well-defined in this setting. We describe the construction in Section 3.1 and use it throughout our arguments.

LEMMA 2.1. *Let ζ and ζ' be two instances of DLAS on a transitive unimodular graph. Let V_t and V'_t be the occupation times of the root by A -particles up to time t . If $\zeta_0(v) \leq \zeta'_0(v)$ a.s. for all v , then $EV_t \leq EV'_t$.*

LEMMA 2.2. *Let $\rho_s^{(r)}$ be the density of A -particles at the root in DLAS with $0 \leq \lambda_B \leq 1$ and $\lambda_A = 1$ at time s with particles beyond graph distance r from the root removed from the initial configuration in $\vec{\mathbb{T}}_{2d}$ or outside of $[-r, r]^d$ removed from the initial configurations in $\mathbb{T}_{2r}^d$ and $\mathbb{Z}^d$.*

For $\vec{\mathbb{T}}_{2d}$ with $d \geq 2$ and $r = \lceil ct \rceil$, it holds for some absolute constant c and all $s \leq t$ that

$$|\rho_s - \rho_s^{(r)}| \leq e^{-t}.$$

For the torus $\mathbb{T}_{2r}^d$ and for $\mathbb{Z}^d$,

$$|\rho_t - \rho_t^{(r)}| \leq C_1 t^d e^{-C_0 r^2/t}$$

for constants $C_0, C_1 > 0$ and all r, t satisfying $1 \leq \sqrt{t} \leq r \leq 2t$.

LEMMA 2.3. *Let ρ_t be the density of A-particles for DLAS on $\mathbb{Z}^d$ with $0 \leq \lambda_B \leq 1$ and $\lambda_A = 1$. Let $\bar{\rho}_t$ be the density of A-particles for DLAS on the torus $\mathbb{T}_{2r}^d$ at time t with $\sqrt{t} \leq r \leq 2t$. There exist $c, C > 0$ that do not depend on t and r such that*

$$|\rho_t - \bar{\rho}_t| \leq C t^d e^{-cr^2/t}$$

for all $t \geq 1$.

In Sections 4.3 and 4.4, we will consider a variant of DLAS that we call the *sequential process run for time t* . Let $\lambda_B = 0$ and fix a time $t \leq \infty$. Assume that DLAS on a given graph has standard initial conditions restricted to a finite subset H of the graph. Place any ordering on the vertices of H , and let it induce an ordering on the A -particles according to their initial positions. With all other particles holding still, let the first A -particle carry out its random walk up to time t or until it hits a B -particle, in which case it and the B -particle are annihilated as usual. Then, in this new landscape of B -particles, run the second A -particle in the same way, and continue until all A -particles have gone. Define the *occupation time* of the root for this process to be the sum of the times spent at the root by each A -particle.

We will show that the sequential process dominates the usual DLAS, in the sense that the occupation time of the root is stochastically larger in the sequential process than in the usual DLAS. For random variables X and Y , we use the notation $X \preceq Y$ to denote that X is *stochastically dominated* by Y in the standard sense that $\mathbb{P}(X \leq t) \geq \mathbb{P}(Y \leq t)$ for all t , or equivalently that there exists a coupling of X and Y so that $X \leq Y$ a.s.

LEMMA 2.4. *Consider DLAS on a transitive unimodular graph G . Let H be a subset of the vertices of G and assume $\lambda_B = 0$. Let V_t be the total occupation time of the root by A -particles up to time $t \leq \infty$ in a DLAS with all B -particles present and only the A -particles initially in H present. For a given ordering of vertices in H , let V'_t be the total occupation time of the root in the sequential process run for time t . Then $V_t \preceq V'_t$.*

Next, imagine dividing the particles in DLAS into two classes and running each as a separate instance of DLAS. Intuitively, the combined particle density in the separated processes should dominate the density in the origin DLAS, since more annihilations will occur in the original process. We prove this for the $\lambda_B = 0$ case.

LEMMA 2.5. *Consider DLAS on a transitive unimodular graph with given initial conditions with no more than one particle per site, and with every particle labeled as positive or negative. Then consider two independent DLAS, one with only the positive particles present and one with only the negative particles present. Let V_t^+ and V_t^- be the total occupation times of the root up to time t in each of these processes. For all $0 < t \leq \infty$, it holds that*

$$V_t \preceq V_t^+ + V_t^-.$$

3. Process construction and proofs of key lemmas. The goal of this section is to establish the lemmas from Section 2. In Section 3.1, we describe the graphical construction of DLAS and a tracer system to track differences between different DLAS from [13]. These tools help prove Lemmas 2.1, 2.2, and 2.3. In Sections 3.2 and 3.3, we introduce two modified tracer constructions that we use to prove Lemma 2.4 and Lemma 2.5, respectively.

We note that [13, 36, 12, 37, 6] developed couplings for DLAS. As mentioned, [13] introduced the tracer construction. A modified tracer system was later used in [6]. A variant of DLAS in which A -particles can selectively ignore certain collisions with B -particles is considered in [36]. The paper also uses a construction of DLAS with $\lambda_B = 0$ where each site has a stack of instructions. This perspective is a special case of a more general Abelian property for the equivalent particle-hole model mentioned earlier [12, Lemma 1]. A path concatenation scheme that uses Young tableaux is used in [37] in a version of internal diffusion-limited aggregation, which can be thought of as a DLAS in which A -particles are released sequentially from a single source. Speaking to the subtlety of DLAS, all of these couplings are different from each other as well as from the modified tracer systems we introduce in Sections 3.2 and 3.3.

3.1. Tracers and the proofs of Lemmas 2.1–2.3. The essential idea in the proofs of Lemmas 2.2 and 2.3 is that when considering the occupation time at a site, particles that begin far away from the site have little effect on the occupation time and can be ignored. To make this precise, we use a construction from [13] in which *tracer particles* track the differences between two coupled DLAS with different initial conditions.

First, we describe the graphical construction for DLAS from [13]. Fix $x \in G$ and assign to the j th particle counted by $\zeta_0(x)$ a discrete simple random walk path $W^{x,j} = (W_k^{x,j})_{k \in \mathbb{N} \cup \{0\}}$ which it follows according to a rate λ_A or λ_B Poisson point process depending on whether j is positive (λ_A) or negative (λ_B). This forms a continuous-time path $S^{x,j}$ called the *putative trajectory* of the particle. We further assign to the j th particle counted by $\zeta_0(x)$ a *braveness* $h^{x,j} \in [0, 1]$ chosen uniformly at random. Assume independence of $\{S^{x,j}, h^{x,j} : x \in G, j \in \mathbb{Z} \setminus \{0\}\}$. Particles follow their putative trajectories $S^{x,j}$. When one or more particles of opposite type occupy a site, the bravest A - and B -particles pairwise mutually annihilate until there are no remaining pairs of opposite type particles at the site. It is shown in [13, Section 6] that for a transitive unimodular graph (such as $\tilde{T}_{2d}$, $\mathbb{T}_r^d$, or $\mathbb{Z}^d$), this graphical model is well defined, and its particle counts at time t form a Markov process with the correct transition rates for DLAS.

The graphical construction allows us to couple DLAS with different initial configurations. Let ζ' be a DLAS defined using the same graphical construction as ζ but with initial conditions ζ'_0 . Since we will only need to add or remove at most one particle from any given site, we will assume that $|\zeta'_0(x) - \zeta_0(x)| \leq 1$ for all $x \in G$. Let $\mathcal{A}^+ = \{x : \zeta'_0(x) - \zeta_0(x) = 1\}$ and $\mathcal{A}^- = \{x : \zeta'_0(x) - \zeta_0(x) = -1\}$. Set $\mathcal{A} = \mathcal{A}^+ \cup \mathcal{A}^-$.

We use a method of *tracer particles* given in [13] to track the difference between ζ and ζ' as they evolve. We now summarize their construction, described in more detail in [13, Section 3.1]. (Though the construction is given within the section of [13] dealing with the case $\lambda_A = \lambda_B = 1$, it works for general λ_A and λ_B .) To describe how tracers work, we follow [13] and start with the case of a single A -particle added at location x , i.e., $\zeta'_0(x) = \zeta_0(x) + 1$ with $\zeta_0(x) \geq 0$. We let $X^x = (X_t^x)_{t \geq 0}$ denote the path of the tracer particle we define now. Initially the tracer follows this A -particle's path in ζ' . Now, suppose this A -particle is annihilated in ζ' . Following the annihilation, there is either a B -particle present in ζ but not ζ' (if no other A -particle was present at the annihilation site), or there remains an extra A -particle present in ζ' but not ζ (if another A -particle was present at the annihilation site). The tracer then follows the extra A -particle in ζ' or the extra B -particle in ζ . When this particle is annihilated, there

will continue to be either an extra A -particle in ζ' or an extra B -particle in ζ , and the tracer transfers itself to this extra particle. This process continues, and $\zeta'_t - \zeta_t = \delta_{X_t^x}$ for all t . If $\lambda_A = \lambda_B = 1$, then X_t^x is a random walk in continuous time. If we only know that $\lambda_A \leq 1$ and $\lambda_B \leq 1$, then X_t^x can be coupled with a rate 1 random walk so that in time t it visits a subset of the sites visited by a rate-1 random walk. Though we described this construction for the case where initially an extra A -particle is initially present in ζ' , it works equally well when an extra B -particle is initially present in ζ , and hence in all cases where $\mathcal{A}^+ = \{x\}$ and $\mathcal{A}^- = \emptyset$. If instead $\mathcal{A}^+ = \emptyset$ and $\mathcal{A}^- = \{x\}$, the same construction yields a tracer that at all times follows an extra B -particle in ζ' or an extra A -particle in ζ .

When $\mathcal{A}$ is not a singleton, the same construction applied to each element of $\mathcal{A}$ yields a collection of tracers. Tracers originating from $\mathcal{A}^+$ are called $\oplus$ -tracers, and those originating from $\mathcal{A}^-$ are called $\ominus$ -tracers. The only complication comes when a $\oplus$ -tracer encounters a $\ominus$ -tracer. This may correspond to two extra particles of the opposite type, either both in ζ or both in ζ' , which annihilate each other. In this case the tracers are left with nothing to track, and the two differences between ζ and ζ' that were tracked by them are no more. It can also correspond to two extra particles of the same type, one in ζ and one in ζ' , both simultaneously annihilated by the same particle in both systems, again eliminating two differences between ζ and ζ' and leaving the tracers nothing to track. In these cases, we say that the two tracers are *wandering* rather than *active* from this point on, but we extend their paths by independent rate- λ_A random walks. (Note that a $\oplus$ - and $\ominus$ -tracer can also encounter each other without the tracers becoming wandering, for example when they track two extra A -particles, one in ζ and one in ζ' , but no B -particle encounters them before they move apart.) We then have

$$(6) \quad \zeta'_t - \zeta_t = \sum_{\sigma \in \{+, -\}} \sum_{y \in \mathcal{A}^\sigma} \sigma \mathbf{1}\{X^y \text{ is active at time } t\} \delta_{X_t^y}.$$

PROOF OF LEMMA 2.1. This lemma follows immediately from Lemma 3 in [13], which asserts that when extra A -particles are added to the graphical construction, the lifespans of all existing A -particles are at least as long as before, and the lifespans of all existing B -particles are no longer than before. (Lemma 3 in [13] is itself an immediate consequence of the tracer construction and (6).) $\square$

We are now ready to prove Lemma 2.2. The idea is to consider DLAS with full and truncated initial conditions and then to bound the difference between the two systems using tracers.

PROOF OF LEMMA 2.2. We first provide the argument for the bidirected tree, then explain how to adapt it to the lattice and torus. Let B_r denote the ball of radius r around the root in $\vec{\mathcal{T}}_{2d}$ and let $r = \lfloor ct \rfloor$ for a constant $c > 1$ to be chosen later. Let ζ denote DLAS on $\vec{\mathcal{T}}_{2d}$, and let ζ' denote the same process with particles beyond graph distance r of the root removed from the initial conditions. Take ζ and ζ' to be coupled by the graphical construction described previously, and for $y \in \vec{\mathcal{T}}_{2d} \setminus B_r$, let $X^y = (X_u^y)_{u \geq 0}$ denote the tracer path originating at y .

Let $\alpha_s(x) = \zeta_s(x) \mathbf{1}\{\zeta_s(x) > 0\}$ and $\alpha'_s(x) = \zeta'_s(x) \mathbf{1}\{\zeta'_s(x) > 0\}$, the number of A -particles present at x at time s in each system. Since $\rho_s - \rho_s^r = \mathbb{E}[\alpha_s(\mathbf{0}) - \alpha'_s(\mathbf{0})]$, and $|\alpha_s(\mathbf{0}) - \alpha'_s(\mathbf{0})|$ is bounded by the number of tracers at $\mathbf{0}$ at time s ,

$$(7) \quad |\rho_s - \rho_s^r| \leq \sum_{y \in \vec{\mathcal{T}}_{2d} \setminus B_r} \mathbb{P}(X_s^y = \mathbf{0}).$$

Now we bound this sum. Consider $y \in \vec{\mathcal{T}}_{2d} \setminus B_r$ such that there is an oriented path in $\vec{\mathcal{T}}_{2d}$ from y to the root of length $u > r$. We claim that for such y and for all $1 \leq s \leq t$,

$$(8) \quad \mathbb{P}(X_s^y = \mathbf{0}) \leq e^{-(u-t)^2/u} d^{-u}.$$

Indeed, as we mentioned when defining tracers, since $\lambda_A, \lambda_B \leq 1$ we can couple X^y with a rate 1 random walk $S = (S_v)_{v \geq 0}$ from y so that X^y visits a subset of the sites traversed by S up to a given time. For S to reach $\mathbf{0}$ by time $s \leq t$, it must make at least u jumps by time s , which occurs with probability at most $e^{-(u-t)^2/u}$ by Lemma A.2 (note that $u \geq t$ since $c > 1$). Also, S must make its first U jumps toward the root if it is ever to reach it, which has probability d^{-u} . Together, this proves (8).

The number of y such that there is an oriented path in $\vec{T}_{2d}$ from y to the root of length u is d^u (see Figure 3), and for y such that there is no oriented path from y to the root, we clearly have $\mathbb{P}(X_s^y = \mathbf{0}) = 0$. Thus, splitting up the sum according to u and recalling $r = \lceil ct \rceil$, it holds for $s \leq t$ that

$$\sum_{y \in \vec{T}_{2d} \setminus B_r} \mathbb{P}(X_s^y = \mathbf{0}) \leq \sum_{u=\lceil ct \rceil}^{\infty} e^{-(u-t)^2/u} \leq e^{2t} \sum_{u=\lceil ct \rceil}^{\infty} e^{-u} \leq e^{-t}$$

for a sufficiently large choice of c , which gives the lemma for the bidirected tree.

For the lattice $\mathbb{Z}^d$, let B_r denote the closed ℓ^∞ -ball of radius r centered at $\mathbf{0}$. Define ζ as DLAS on $\mathbb{Z}^d$ and ζ' as the same with particles outside of B_r removed from the initial conditions. As before,

$$|\rho_s - \rho_s^r| \leq \sum_{y \in \mathbb{Z}^d \setminus B_r} \mathbb{P}(X_s^y = \mathbf{0}).$$

The argument below holds whenever $1 \leq r \leq 2t$, but the resulting bound will only be meaningful when $r \geq \sqrt{t}$.

Consider $y \in \mathbb{Z}^d \setminus B_r$. As in the $\vec{T}_{2d}$ case, we bound $\mathbb{P}(X_s^y = \mathbf{0})$ by the probability of a rate 1 random walk $S = (s_v)_{v \geq 0}$ from y hitting $\mathbf{0}$ before time s . Using a crude bound here, let $u = \|y\|_\infty > r$ and choose a coordinate of y whose absolute value is u . The projection of S onto this component is a rate $1/d$ 1-dimensional random walk, and its probability of reaching 0 by time $s \leq t$ is at most $e^{-u^2/4t}$ by Lemma A.1. Using this bound and letting $b_u = |\{x \in \mathbb{Z}^d : \|x\|_\infty = u\}| \leq C_d u^{d-1}$, the expected number of tracers started between distance $r+1$ and $2t$ from $\mathbf{0}$ that reach $\mathbf{0}$ by time t is at most (recall $r \leq 2t$)

$$(9) \quad \sum_{y \in B_{2t} \setminus B_r} \mathbb{P}(X_s^y = \mathbf{0}) \leq \sum_{u=r+1}^{2t} b_u e^{-u^2/4t} \leq \sum_{u=r+1}^{2t} C_d u^{d-1} e^{-u^2/4t} = O(t^{d-1} e^{-r^2/4t}).$$

The asymptotic bound is immediate when $d = 1$ and for $d \geq 2$ is obtained by comparing the last sum in (9) to the integral

$$(10) \quad \begin{aligned} \int_r^{2t} u^{d-1} e^{-u^2/4t} &\leq (2t)^{d-2} \int_r^{2t} u e^{-u^2/4t} du \\ &= C' t^{d-1} \int_{r^2/4t}^t e^{-w} dw \leq C' t^{d-1} e^{-r^2/4t}. \end{aligned}$$

Now, we use Lemma A.2 to bound

$$(11) \quad \sum_{y \in \mathbb{Z}^d \setminus B_{2t}} \mathbb{P}(X_s^y = \mathbf{0}) \leq \sum_{k=1}^{\infty} b_{2t+k} \exp\left(\frac{-(t+k)^2}{2(2t+k)}\right) = O(t^d e^{-c't}).$$

The asymptotic claim follows from a similar approach as at (10). Combining (9) and (11) with (7) gives the claimed bound. A similar argument gives the analogous bound for $\mathbb{T}_{2r}^d$. $\square$

We now prove that DLAS on $\mathbb{T}_{2r}^d$ is comparable to DLAS on $\mathbb{Z}^d$. We start with an outline of the argument. First we couple the particle types and random walk paths at corresponding sites of the torus and lattice. Letting F be the event that a particle started within distance r of the root interacts with the boundary, we show that $\mathbb{P}(F)$ is small. We then show that the density of particles at the root of the torus and of the lattice does not change much when all particles beyond distance r are removed. With this removal, the processes on the torus and lattice are identical when F^c occurs. And on the event F , the density of particles at the root can be easily controlled by comparing to DLAS systems with only A -particles present. Since $\mathbb{P}(F)$ is small, we obtain a good estimate on $|\rho_t - \bar{\rho}_t|$.

PROOF OF LEMMA 2.3. Identify each site $\bar{x} \in \mathbb{T}_{2r}^d$ with $x \in \mathbb{Z}^d \cap (-2r, 2r]^d$ in the canonical way that comes from viewing $\mathbb{T}_{2r}^d$ as a quotient space on the lattice with points in $(-2r, 2r]^d$ as representatives of each equivalence class. We now couple DLAS on $\mathbb{Z}^d$ and on $\mathbb{T}_{2r}^d$ using the graphical construction from [13]. First, couple the initial configurations in $\mathbb{T}_{2r}^d$ and $\mathbb{Z}^d \cap (-2r, 2r]^d$ to be the same. Using the natural identification of points outside of $(-2r, 2r]^d$ to the equivalence class representative in $\mathbb{T}_{2r}^d$, couple the random walk instructions at $\bar{x} \in \mathbb{T}_{2r}^d$ and $x \in \mathbb{Z}^d \cap (-2r, 2r]^d$ to be the same. For $x \in \mathbb{Z}^d \setminus (-2r, 2r]^d$, let the initial configuration and instructions be generated independently. We will refer to sites in $\mathbb{T}_{2r}^d$ with some coordinate entry equal to $2r$ as *boundary sites* of $\mathbb{T}_{2r}^d$.

Let $B_r = \{\bar{x} \in \mathbb{T}_{2r}^d : \|x\|_\infty \leq r\}$. Let $D_{\bar{x}}$ be the event that in time t the random walk instructions for the particle started at $\bar{x} \in B_r$ reaches a boundary site of $\mathbb{T}_{2r}^d$. As $\bar{x} \in B_r$, the distance from $\bar{x}$ to a boundary site is at least r . Hence, $\mathbb{P}(D_{\bar{x}})$ is bounded by the probability that a rate 1 simple random walk has displacement at least r by time t . Each coordinate of the d -dimensional simple random walk on $\mathbb{Z}^d$ forms a rate $1/d < 1$ simple random walk on $\mathbb{Z}$. Let $F = \cup_{\bar{x} \in B_r} D_{\bar{x}}$ be the event that some particle started inside of B_r reaches a boundary site of the torus by time t . Lemma A.1 and a union bound over the initial locations $\bar{x} \in B_r$ and d coordinates give

$$(12) \quad \mathbb{P}(F) \leq d|B_r|e^{-r^2/4t} \leq d(2r)^d e^{-r^2/4t} \leq d(4t)^d e^{-r^2/4t}.$$

Note that Lemma A.1 still applies to the rate $1/d$ random walk since the maximum of this random walk is dominated by that of a rate 1 random walk.

Let ρ_t^* and $\bar{\rho}_t^*$ be the expected density of particles at $\mathbf{0}$ at time t for the lattice and torus, respectively, with all particles initially beyond distance r deleted (in the ℓ^∞ -norm). By Lemma 2.2, there are constants $C, c > 0$ for which we have

$$(13) \quad |\rho_t - \rho_t^*|, |\bar{\rho}_t - \bar{\rho}_t^*| \leq Ct^d e^{-cr^2/t}.$$

Hence, it suffices to compare ρ_t^* and $\bar{\rho}_t^*$.

Let Z and $\bar{Z}$ be the number of A -particles at the origin at time t on the lattice and torus, respectively, with the particles beyond distance r deleted from the initial configuration. We claim that

$$Z\mathbf{1}\{F^c\} = \bar{Z}\mathbf{1}\{F^c\}$$

since, on the event F^c , the random walk paths are identical for corresponding particles in the two coupled processes. It follows that

$$|\rho_t^* - \bar{\rho}_t^*| = |\mathbb{E}[(Z - \bar{Z})\mathbf{1}\{F\}]|.$$

The Cauchy-Schwartz inequality and the simple bound $\mathbb{E}[(Z - \bar{Z})^2] \leq \mathbb{E}[Z^2] + \mathbb{E}[\bar{Z}^2]$ give

$$(14) \quad |\mathbb{E}[(Z - \bar{Z})\mathbf{1}\{F\}]| \leq \sqrt{\mathbb{E}[Z^2] + \mathbb{E}[\bar{Z}^2]} \sqrt{\mathbb{P}(F)}.$$

It follows from [13, Lemma 3] that both Z and $\bar{Z}$ are dominated by the counts of A -particles at the root in systems with an A -particle initially at every site. Call these dominating counts on the lattice and torus Z' and $\bar{Z}'$, respectively. By symmetry of the underlying graphs, $\mathbb{E}[Z'] = 1 = \mathbb{E}[\bar{Z}']$ for all $t \geq 0$. Using this fact and expressing Z' and $\bar{Z}'$ as sums of independent indicators for whether or not the particle started at each site is at $\mathbf{0}$ at time t , it is straightforward to prove that $\mathbb{E}[(Z')^2], \mathbb{E}[(\bar{Z}')^2] \leq 2$. In full detail, letting p_x be the probability a particle started at x is at $\mathbf{0}$ at time t , expanding $\mathbb{E}[(Z')^2]$ and $(\mathbb{E}[Z'])^2$ gives

$$\mathbb{E}[(Z')^2] = 2 \sum_{x,y \in \mathbb{Z}^d, x \neq y} p_x p_y + \sum_{x \in \mathbb{Z}^d} p_x \leq (\mathbb{E}[Z'])^2 + \mathbb{E}[Z'] = 2.$$

Proceeding with similar reasoning also gives that $\mathbb{E}[(\bar{Z}')^2] \leq 2$.

It follows from (14) that

$$(15) \quad |\rho_t^* - \bar{\rho}_t^*| \leq 2\sqrt{\mathbb{P}(F)}.$$

Using (12), we have $\mathbb{P}(F) \leq C_3 t^d e^{-c_4 r^2/t}$ with $c_4, C_3 > 0$. Applying this to (15) and then using (13) with the triangle inequality gives

$$|\rho_t - \bar{\rho}_t| \leq C' t^d e^{-c' r^2/t}$$

for some constants $c', C' > 0$. □

3.2. A variation on tracers and the proof of Lemma 2.4. We now give another construction of DLAS designed to relate it to the sequential process defined in Section 2. It resembles the dragged tracer construction from [13, Section 4.1] but is not quite the same (see Remark 3.3).

We assume now that $\lambda_A = 1$ and $\lambda_B = 0$. We will define a particle system with three types of particles: B -particles, A -tracers, and B -tracers. As we will see in Proposition 3.2, if we view the A -tracers as A -particles and ignore the B -tracers altogether, the resulting system will have the law of DLAS.

We now describe this process, which we call the *A/B-tracer system*. At time 0, the system consists of possibly infinitely many B -particles, no more than one per site, as well as a finite number of particles we call A -tracers, which we number $1, \dots, n$. The k th A -tracer is assigned a rate-1 random walk $S^k = (S_t^k)_{t \geq 0}$ as its *putative trajectory*. The A -tracers initially follow these trajectories, while B -particles do not move. When an A -tracer jumps onto a B -particle, the B -particle is annihilated and the A -tracer becomes a B -tracer and halts. If the j th A -tracer jumps onto the k th B -tracer with $j < k$, then the A -tracer becomes a B -tracer and halts, while the B -tracer becomes an A -tracer and resumes following the path S^k from where it left off when it became a B -tracer; if $j > k$ then no interaction occurs.

We claim that under these dynamics, there is no need to assign a braveness to each particle because particles can never achieve a configuration where it is ambiguous which particles should react. Under these interaction rules and our assumptions on the initial configuration, there is at most one B -particle or B -tracer on a site at all times. Since A -tracers move in continuous time, almost surely only a single A -tracer jumps onto a B -particle or B -tracer at a time. And while an A -tracer may jump onto a site that contains A -tracers and a B -tracer, the A -tracers already present will have indices greater than the B -tracer's, and only the newly arrived A -tracer may interact with the B -tracer. Also, since the system contains only finitely many moving particles, there is no question that the construction is well defined. We record two immediate observations for later reference:

LEMMA 3.1. *In the A/B-tracer system:*

- (a) a site that does not start with a B -particle will never contain B -particles or B -tracers;
- (b) a site that initially contained a B -particle still contains the B -particle if it has not been visited by A -particles yet; otherwise it contains a B -tracer whose index is the lowest of all tracers that have visited the site up to that point.

As we said before, the A/B -tracer system gives an alternative construction of DLAS:

PROPOSITION 3.2. *Define $\zeta_t(x)$ as the number of A -tracers minus the number of B -particles at x at time t in the particle system defined above. Then ζ is a DLAS with $\lambda_A = 1$ and $\lambda_B = 0$.*

PROOF. Simply observe that ζ_t is a Markov process with the same dynamics as DLAS. $\square$

REMARK 3.3. *The dragged tracer construction of [13, Section 4.1] is like the system of tracers (also from [13]) described in Section 3.1, except that tracers have their own pre-specified paths, rather than simply following the prespecified paths of the particles they are tracing. The A/B -tracer system resembles the dragged tracer construction that would result from adding a collection of A -particles to a DLAS that has only B -particles. But the two constructions differ in that in our system, an A -tracer can interact with a B -tracer, whereas two $\oplus$ -tracers in [13] never interact.*

The A/B -tracer system could also be defined when $\lambda_B > 0$. We do not do so here because we do not need it, and the possibility of a tracer jumping onto a site containing multiple particles and tracers of the opposite type adds some technical difficulty. A closely related construction appears in [6]. The system defined there has exactly two tracers and also A -particles that interact with the tracers by similar rules. It still has $\lambda_B = 0$, but it allows multiple B -particles to start on the same site and for two A -tracers to jump on a site simultaneously in discrete time.

We defined ζ in terms of the A/B -tracer system by having it count A -tracers as A -particles and ignore B -tracers. If we instead define ζ^k to count tracers $1, \dots, k$ as A -particles when in state A , to count tracers $k+1, \dots, n$ as B -particles when in state B , and to ignore the tracers in other states, then ζ^k is a DLAS that initially has k A -particles. This gives a coupling of the sequence of systems resulting from successively adding A -particle one at a time. We do not need this for this paper, though it is used in [6].

We now modify the A/B -tracer system by killing each tracer particle when its elapsed time spent in state A reaches t . We call the resulting system the t -killed A/B -tracer system. In more detail, at a given time s , each tracer particle in the A/B -tracer system has had a finite number of interactions that cause it to switch from an A -tracer to a B -tracer or vice versa. Let $L_k(s)$ denote the combined length of the time periods up to s that the k th tracer has spent as an A -tracer, so that its position at time s is $S_{L_k(s)}^k$. At time $U_k := \inf\{s : L_k(s) \geq t\}$, the k th tracer is killed and hence removed from the system. Observe that the k th tracer may enter state B and never leave it before spending time t in state A , in which case $U_k = \infty$. We define $T_k = t$ if $U_k < \infty$ and $T_k = \max\{L_k(s) : s \geq 0\}$ if $U_k = \infty$. Thus the k th tracer traces out the path $(S_s^k)_{0 \leq s \leq T_k}$ over the lifetime of the process, pausing whenever it is in state B . If $t = \infty$, then we define $U_k = \infty$ and $T_k = \sup\{L_k(s) : s \geq 0\}$, with $T_k = \infty$ if $L_k(s)$ is unbounded. The resulting process is the original A/B -tracer system and as in the $t < \infty$ case, the k th tracer follows the path $(S_s^k)_{0 \leq s \leq T_k}$ over its lifetime.

PROOF OF LEMMA 2.4. Consider the sequential process run for time $t \leq \infty$. For now, take H to be finite and let n be the number of A -particles in the set H . For $k \in \{1, \dots, n\}$, let $S^k = (S_s^k)_{s \geq 0}$ denote the random walk followed by the k th A -particle. Let T'_k be the length of time spent walking by the k th A -particle in the sequential process before annihilation, with $T'_k = t$ if it is never annihilated. Thus the path of the k th particle in the sequential process is $(S_s^k)_{0 \leq s \leq T'_k}$.

Now, we consider the t -killed A/B -tracer system with initial conditions corresponding to the sequential process, that is, with A -tracers initially at locations in H containing A -particles and with B -particles initially at all sites in G with B -particles. We give the tracers the same ordering as in the sequential system, and we couple the sequential process with the t -killed A/B -tracer system by using the same collection of random walks $\{S^1, \dots, S^n\}$ for both processes. Recall that T_k was defined so that $(S_s^k)_{0 \leq s \leq T_k}$ is the path traced out by the k th tracer in this system.

We claim that $T_k = T'_k$ for all k , and hence that the paths traced out in the sequential process and t -killed A/B -tracer systems are the same. The rest of the lemma will follow easily once this claim is proved, since the original and t -killed A/B -tracer systems do not differ until at least time t .

We prove the claim by induction on k . To prove that $T_1 = T'_1$, we observe that the first A -tracer enters state B on colliding with any B -particle or B -tracer. Since every site that starts with a B -particle contains a B -particle or a B -tracer at all times, the first A -tracer becomes a B -tracer exactly when it visits a site that originally contained a B -particle. Since it has the lowest index, it can never turn back to an A -tracer. Thus T_1 is the first time that S^1 visits a site that originally contained a B -particle, or $T_1 = t$ if S^1 does not visit such a site by time t . This is exactly the description of T'_1 as well, proving that $T_1 = T'_1$.

Now, suppose that $T_j = T'_j$ for $j = 1, \dots, k-1$, and we will show that $T_k = T'_k$. First, observe that T_k and T'_k must either be jump times of the walk S^k or be equal to t . For T_k , this is because if $T_k < t$, then the k th tracer entered state B at some time u satisfying $T_k = L_k(u)$ and never reentered state A . Since an A -tracer enters state B only upon jumping, T_k is a jump time of S^k . Similarly, for T'_k it is because a particle in the sequential process is annihilated only at a jump time of S^k .

Thus, to show $T_k = T'_k$ it will suffice to show that for jump times r of S^k satisfying $r < t$, we have $T_k = r$ if and only if $T'_k = r$. We can also assume that the first visit by S^k to S_r^k occurs at time r , since in the t -killed A/B -process the k th A -tracer can only enter state B on its first visit to a site, and in the sequential process an A -particle can only be annihilated on its first visit to a site. For such a jump time $r < t$ with $x := S_r^k$, we claim that $T_k = r$ if and only if the following conditions hold:

- (i) $T_k \geq r$;
- (ii) there is initially a B -particle at x in the t -killed A/B -tracer system; and
- (iii) for $j \in \{1, \dots, k-1\}$, the path $(S_u^j)_{0 \leq u \leq T_j}$ does not contain x .

Suppose the above conditions hold. Since $T_k \geq r$, the k th tracer survives long enough to jump to site x at some time s with $r = L_k(s)$. Since there is initially a B -particle at x and the tracers of index smaller than k do not visit x by (iii), there remains a B -particle or B -tracer of index greater than k by Lemma 3.1(b) that the k th tracer will jump onto. The k th tracer then enters state B and never leaves, since no B -particle or B -tracer of index smaller than k visits x . Hence $T_k = r$. Conversely, suppose $T_k = r$. Then the k th tracer jumps onto site x at some time s with $r = L_k(s)$, enters state B , and never leaves. Clearly (i) holds. By Lemma 3.1(b), condition (ii) holds, and no tracer of index less than k visits x before time s . And since the k th tracer never leaves state B , no tracer of index less than k visits x after time s either, showing that (iii) holds.

Next, observe that for a jump time r of S^k with $x := S_r^k$ and with S^k assumed not to visit x until time r , we have $T_k' = r$ if and only if the following conditions hold for the k th A -particle in the sequential process:

- (i') $T_k' \geq r$, that is, the particle has not been annihilated before time r ;
- (ii') there is initially a B -particle at x in the sequential process; and
- (iii') none of the first $k - 1$ A -particles in the sequential process visited x , so for $j \in \{1, \dots, k - 1\}$ the path $(S_u^j)_{0 \leq u \leq T_j'}$ does not contain x .

By our coupling, (ii) holds if and only if (ii') does. By the inductive hypothesis, (iii) holds if and only if (iii') does. And by doing induction on r starting with the earliest jump time, we can assume that (i) holds if and only if (i') holds. This proves that $T_k = T_k'$, advancing the induction and proving that $T_k = T_k'$ for all k .

Let $V_s^{(t)}$ be the occupation time of the root by A -tracers up to time s in the t -killed A/B -tracer system. Then

$$\begin{aligned} V_t^{(t)} &= \sum_{k=1}^n \int_0^t \mathbf{1}\{\text{the } k\text{th } t\text{-killed tracer is in state } A \text{ at } 0 \text{ at time } s\} ds \\ &\leq \lim_{s \rightarrow \infty} V_s^{(t)} = \sum_{k=1}^n \int_0^{T_k} \mathbf{1}\{S_s^k = 0\} ds. \end{aligned}$$

Since the t -killed A/B -tracer system matches the A/B -tracer system at least up to time t , by Proposition 3.2 the random variable $V_t^{(t)}$ is distributed as the occupation time V_t of the root in DLAS with initial conditions as in the statement of the lemma. Let V_t' be the occupation time of the root in the sequential process, i.e.,

$$V_t' = \sum_{k=1}^n \int_0^{T_k'} \mathbf{1}\{S_s^k = 0\} ds.$$

Since $T_k = T_k'$ for all k , this proves that $V_t \preceq V_t'$.

Now, suppose that H is infinite. Let H_n consist of the first n vertices in H in the given ordering, and let $V_{t,n}$ and $V_{t,n}'$ be the occupation times of the root in DLAS and the sequential process, respectively, when A -particles outside of H_n are removed from the initial configuration. By the case of this lemma already proven, $V_{t,n} \preceq V_{t,n}'$. When $V_{t,n}$ is defined in terms of the graphical construction from [13], it increases in n by [13, Lemma 3] and converges almost surely to V_t . By definition of the sequential process, $V_{t,n}'$ converges upwards to V_t' . Thus $V_{t,n} \rightarrow V_t$ and $V_{t,n}' \rightarrow V_t'$ in law as $n \rightarrow \infty$ and $V_{t,n} \preceq V_{t,n}'$ for all n , which together prove that $V_t \preceq V_t'$. $\square$

3.3. Polarized construction of DLAS and proof of Lemma 2.5. We give one last construction of DLAS that we call the *polarized system*. As in the previous section's construction, the system has tracers that can be either in state A or state B —we call them A -tracers or B -tracers depending on their current state—as well as B -particles. We assume that there are only finitely many tracers in the initial configuration. Each tracer starts in state A and is given a rate-1 random walk as its putative trajectory, and it follows it while in state A . The B -tracers and B -particles are immobile.

Each tracer and B -particle in the polarized system is assigned a *polarity*, either *positive* or *negative*, as part of the initial configuration. In the following interaction rules, these polarities play a role similar to the indices in the A/B -tracer system:

- (a) If an A -tracer jumps onto a B -particle, then the B -particle is annihilated and the A -tracer enters state B , regardless of their polarities. (The same occurs in the A/B -tracer system.)

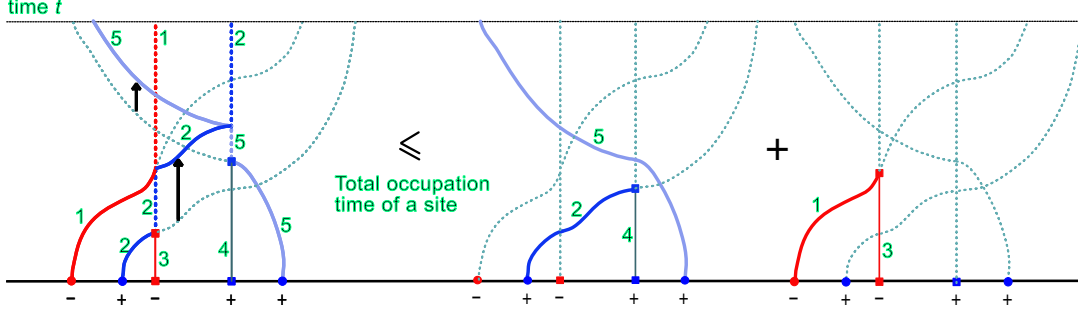


FIG 2. A realization of the polarized system (left) with three initial A -tracers and two B -particles, and the corresponding behavior of the positive (middle) and negative (right) DLAS processes. Dots and squares represent initial A -tracers and B -particles, respectively. In the left system, solid curves are trajectories of A -particles, solid vertical lines trajectories of B -particles, dotted curves are putative trajectories of A -tracers, and dotted vertical lines are trajectories of B -tracers. There are five trajectories in the left system and they are marked by numbers from 1 through 5. In the middle and the right systems, solid curves and vertical lines represent trajectories of A - and B -particles, respectively. Blue and red indicate that the original source of the path was the putative trajectory of a positive or negative particle, respectively. Two different shades of blue are used to disambiguate the two putative trajectories of positive A -tracers. Green dotted curves indicate portions of putative trajectories not used in their respective system (but in the polarized system on left some are used after a time shift, indicated by black upward arrows). At each time, every putative trajectory is traversed no further in the polarized system (left) than in each of the positive (middle) and negative (right) DLAS, so the occupation time of the origin by a given time t in the polarized system is at most the sum in the corresponding positive and negative DLAS.

- (b) If an A -tracer jumps onto a B -tracer of opposite polarity at site v , then whichever tracer matches the polarity of the B -particle initially at v enters (or remains in) state B . The other tracer enters (or remains in) state A and continues along its putative trajectory.
- (c) If an A -tracer jumps onto a B -tracer of the same polarity at site v , then whichever tracer's putative trajectory contains v at the earliest time enters (or remains in) state B . The other tracer enters (or remains in) state A and continues along its putative trajectory.

As with the A/B -tracer system, under these rules there is exactly one B -particle or B -tracer at all times on each site that initially holds a B -particle. When an A -tracer jumps onto a site with a B -tracer, the site contains no other A -tracers or it contains A -tracers that do not interact with the B -tracer, and there is no ambiguity about which A -tracer can interact with the B -tracer. By our assumption of having only finitely many tracers, the system is a Markov chain on a countable state space and there is no question as to its existence.

Now, we define the *positive* and *negative* DLAS via the graphical construction from [13] and then couple these DLAS with the polarized system. We define the *positive* DLAS as follows. For every site in the polarized system that starts with a positive A -tracer, the positive DLAS starts with an A -particle. For every site in the polarized system that starts with a positive B -particle, the positive DLAS starts with a B -particle. At all other sites the positive DLAS initially has no particles. Each A -particle in the positive DLAS is given the same putative trajectory as the corresponding A -tracer in the polarized system. Since $\lambda_A = 1$ and $\lambda_B = 0$ for this DLAS and we start with at most one B -particle per site, we do not need to assign a braveness to our particles. The *negative* DLAS is defined the same way, but its initial configuration matches up with the negative particles in the polarized system. The positive and the negative DLAS are then two DLAS, independent conditional on their initial configurations, both coupled with the polarized system.

Our first claim is that if we ignore B -tracers in the polarized system, we obtain DLAS with the initial conditions given by the A -tracers and B -particles (recall that in the polarized

system all tracers start in state A). The statement and proof are essentially the same as for Proposition 3.2.

PROPOSITION 3.4. *Let $\zeta_t(v)$ be the count of A -tracers minus the count of B -particles on site v at time t in the polarized system. Then ζ_t is DLAS with $\lambda_A = 1$ and $\lambda_B = 0$.*

PROOF. When we are blinded to the polarity of the particles and to B -tracers, the interaction between A - and B -tracers has no effect on ζ_t except replacing a particle's future random walk trajectory with an independent one. And when an A -tracer moves onto a B -particle, the A -tracer enters state B and the B -particle is destroyed, and neither is counted anymore by ζ_t . As in Proposition 3.2, the process ζ_t is then a Markov chain with the same interaction rules as DLAS with $\lambda_A = 1$ and $\lambda_B = 0$. $\square$

PROOF OF LEMMA 2.5. First we suppose that the given DLAS contains only finitely many A -particles. Form a polarized system by assigning polarities according to the signs of particles in the given DLAS. This polarized system is coupled with the positive and negative DLAS, made up respectively of the positive and negative particles only from the given DLAS. By Proposition 3.4, the polarized system provides a construction of the given DLAS. Thus we obtain a coupling of the given DLAS together with the two DLAS, which are independent conditional on their initial configurations, consisting of the negative and positive particles.

Let $S^v = (S_t^v)_{t \geq 0}$ denote the putative trajectory of the A -tracer initially at v , for all sites v initially containing A -tracers in the polarized system. With similar notation as in the A/B -tracer system, we define $L_v(t)$ as the combined time spent by the tracer initially at v in state A up to time t , so that its position at time t is $S_{L_v(t)}^v$.

Suppose that v initially contains an A -tracer in the polarized system (or equivalently that v initially contains an A -particle in the given DLAS). Let $T_v \leq \infty$ denote the time of annihilation of the A -particle in the corresponding negative or positive DLAS, depending on the tracer's polarity. We will argue that $L_v(t) \leq T_v$ for all $t \geq 0$, that is, a tracer in the polarized process never travels farther along its putative trajectory than the corresponding A -particle does in the positive or negative DLAS.

To prove this claim, we must show that if the particle starting at v in the positive or negative DLAS is annihilated upon visiting u , then in the polarized process the tracer starting at v will permanently enter state B upon visiting u . The idea is that interaction rules (b) and (c) of the polarized process ensure that this tracer has priority over all other tracers to be in state B . To make this argument precise, we consider the first time that it fails and derive a contradiction. That is, suppose that our claim is false, and let t be the infimum of all times s such that there exists a vertex v with $L_v(s) > T_v$. Since the system has only finitely many A -particles and $T_v > 0$ for all v , we have $t > 0$. Let v be a site such that $L_v(t + \epsilon) > T_v$ for all $\epsilon > 0$, and let $u = S_{T_v}^v$. As a shorthand, we write α to refer to the tracer in the polarized process initially at site v . For the sake of exposition assume that α is positive. We write α^+ as a shorthand for the particle in the positive DLAS initially at v , which has the same putative trajectory as α . To summarize the set-up, particle α^+ is annihilated upon jumping onto site u at time T_v , but tracer α is on u at time t with $L_v(t) = T_v$ in state A and is about to continue on its putative trajectory past the point that α^+ was annihilated. Our job is to find a contradiction.

When α jumps onto site u , at some time $t' \leq t$, we claim that it enters state B . Since α^+ is annihilated at u , we know that initially there is a positive B -particle at u in the positive DLAS and in the polarized process. By the dynamics of the system, there is a B -particle or B -tracer at u at all times in the polarized process. If the particle on u at time t' is a B -particle, then α enters state B . If the particle on u at time t' is a negative B -tracer, then α enters state A by the

rules of the system. And the particle on u cannot be a positive B -tracer: If it were, then some other positive tracer particle has visited u before time t' . By our minimality assumption for t , this particle traverses no more of its putative trajectory than its corresponding A -particle in the positive DLAS. But this is a contradiction, since then this A -particle in the positive DLAS visits u before time t' , which would mean that $\mathfrak{a}^+$ was not the first A -particle in the positive DLAS to visit u .

Now, we argue that after entering state B at time t' , particle $\mathfrak{a}$ never returns to state A . By the dynamics of the polarized process, it can only return to state A if it is visited by another positive A -tracer whose putative trajectory contains u earlier than time $L_v(t) = T_v \leq t$. But again, by our minimality assumption for t , this would imply that some A -particle in the positive DLAS traverses this putative trajectory up to this visit to u , which contradicts $\mathfrak{a}^+$ being the first A -particle in the positive DLAS to visit u .

Thus we have shown that $\mathfrak{a}$ enters state B at some time $t' \leq t$ and never returns to state A . This contradicts our assumption that $\mathfrak{a}$ is in state A at time t , completing our proof that $L_t(v) \leq T_v$ for all v .

The lemma itself now follows easily. Let v be a site initially containing an A -tracer in the polarized process. Since we have shown that $L_t(v) \leq T_v$ and clearly $L_t(v) \leq t$, a tracer in the polarized process traverses at most length $\min(t, T_v)$ of its putative trajectory in time t . On the other hand, the A -particle initially at v in the positive or negative DLAS traverses exactly length $\max(t, T_v)$ of its putative trajectory in time t . Thus, the time spent at the root by any A -tracer in the polarized process up to time t is at most the time spent by the corresponding A -particle in the negative or positive DLAS, establishing that $V_t \leq V_t^+ + V_t^-$ under this coupling.

Finally, we extend the theorem to initial configurations with infinitely many A -particles by a similar argument as in the proof of Lemma 2.4. Enumerate the vertices of the graph in any way, and let $V_{t,k}$ denote the occupation time of the root up to time t in the given DLAS with all A -particles beyond the first k vertices of the graph removed from the initial configuration. Let $V_{t,k}^-$ and $V_{t,k}^+$ be the analogous quantities for the negative and positive DLAS. By monotonicity of the graphical construction [13, Lemma 3], these random variables converge in law to V_t , V_t^- , and V_t^+ , respectively, as $k \rightarrow \infty$. By the case of the lemma already proven, $V_{t,k} \preceq V_{t,k}^- + V_{t,k}^+$, which together with the convergence proves that $V_t \preceq V_t^- + V_t^+$. And last, a similar limit argument proves the lemma in the $t = \infty$ case. $\square$

4. Proof of main results on the integer lattice. In this section we prove our main results for DLAS on the integer lattice, which are stated in Theorems 1.1, 1.2, 1.3, and 1.4. As defined in Section 1.2, we use $D(H_0)$ to denote the number of A -particles minus the number of B -particles on a finite subgraph H_0 in the initial configuration of a given instance of DLAS.

4.1. *A lower bound on ρ_t for DLAS on $\mathbb{Z}^d$.* We first derive Theorem 1.1 from Lemma 2.3.

PROOF OF THEOREM 1.1. Let $r = \lceil C\sqrt{t \log t} \rceil$, and let $\bar{\rho}_t$ be the density of A -particles at time t on $\mathbb{T}_r^d$ as in Lemma 2.3. By this lemma, we may choose C large enough so that

$$|\rho_t - \bar{\rho}_t| \lesssim (t \log t)^{-d/4}.$$

It then suffices to show that

$$\bar{\rho}_t \gtrsim (t \log t)^{-d/4}.$$

Let $D = D(\mathbb{T}_r^d)$ denote the difference between the initial number of A - and B -particles on $\mathbb{T}_r^d$. We can express D as the sum of $|T_r^d|$ i.i.d. random variables taking value in $\{-1, 1\}$ with

mean 0. The central limit theorem and the fact that $|\mathbb{T}_r^d| \asymp (t \log t)^{d/2}$ guarantee that there is a constant $c > 0$ so that

$$\inf_{t \geq 0} \mathbb{P} \left(D > (t \log t)^{d/4} \right) > c.$$

Notice that on the event $D > (t \log t)^{d/4}$, there are at least $(t \log t)^{d/4}$ A -particles on the torus that survive forever. Hence the expected number of A -particles in the system at time t is at least $c(t \log t)^{d/4}$. By translation invariance,

$$\bar{\rho}_t \geq \frac{c(t \log t)^{d/4}}{|\mathbb{T}_R^d|} \geq c'(t \log t)^{-d/4}.$$

This completes the proof. $\square$

4.2. A lower bound on the critical exponent in low dimension. The idea for the proof of Theorem 1.3 was shared with us by Michael Damron. To give a lower bound on the critical exponent for $\mathbb{E}_p V$, we replicate the proof of Theorem 1.1 for $p < 1/2$ using a more refined estimate in place of the central limit theorem.

LEMMA 4.1. *Let $D = D(\mathbb{T}_r^d)$, the discrepancy of the initial configuration of DLAS on the torus. Then there exists an absolute constant $c_1 > 0$ so that the following implication holds for all $1/4 \leq p < 1/2$:*

$$2r \leq \left(\frac{c_1}{1-2p} \right)^{2/d} \implies \mathbb{P}_p(D \geq c_1(2r)^{d/2}) \geq c_1.$$

PROOF. Let $n = (2r)^d$, the volume of $\mathbb{T}_r^d$. The random variable D is the sum of n independent random variables taking values 1 and -1 with probability p and $1-p$, respectively. Thus $\mathbb{E}D = -n(1-2p)$ and $\sigma^2 := \text{Var } D = 4np(1-p)$. The idea of the proof is that the condition $2r \leq (c/(1-2p))^{2/d}$ ensures that D has mean and standard deviation on the order of $\sqrt{n} = (2r)^{d/2}$, and hence $\mathbb{P}_p(D \geq c(2r)^{d/2})$ is bounded from below. We make this precise now using a quantitative CLT.

Let $c > 0$ be a small constant to be determined later. Suppose that $2r \leq (c/(1-2p))^{2/d}$. It follows that $1-2p \leq c/\sqrt{n}$. From our assumption that $1/4 \leq p < 1/2$, we have $\sigma^2 \geq 3n/4$. Thus

$$\mathbb{E}D + \frac{4c\sigma}{\sqrt{3}} \geq c\sqrt{n} = c(2r)^{d/2}.$$

With Z a standard normal random variable, we obtain

$$\mathbb{P}_p(D \geq c(2r)^{d/2}) \geq \mathbb{P}_p \left(D \geq \mathbb{E}D + \frac{4c\sigma}{\sqrt{3}} \right) \geq \mathbb{P} \left(Z \geq 4c/\sqrt{3} \right) - \frac{C}{\sqrt{n}}$$

for some absolute constant C by the Berry–Esseen CLT. Take $c = c'$ sufficiently small so that $\mathbb{P}(Z \geq 4c'/\sqrt{3}) \geq 1/4$. Then the lower bound above is at least $1/8$ for all $n \geq (8C)^2$ since $C/\sqrt{n} \leq 1/8$. By considering the finite number of small values of n not covered by the previous bound, we get $\mathbb{P}_p(D \geq c'(2r)^{d/2}) \geq c''$ for all $n \geq 1$ for some constant $c'' > 0$. Then the desired lower bound in the assertion holds by letting $c_1 = \min(c', c'')$. $\square$

PROOF OF THEOREM 1.3. Let $\epsilon = 1 - 2p$. Given ϵ , we define

$$r = r(\epsilon) = \left\lfloor \frac{1}{2} (c_1/\epsilon)^{-2/d} \right\rfloor,$$

where c_1 is the constant in Lemma 4.1. Thus, this lemma shows that if X is the number of A -particles that survive forever for DLAS on $\mathbb{T}_r^d$, then $\mathbb{E}_p X \geq c_1^2 (2r)^{d/2}$. Let $\bar{\rho}_t$ be the density of A -particles in DLAS on $\mathbb{T}_r^d$. By translation invariance, for all $s \geq 0$ and $p \in [1/4, 1/2)$,

$$(16) \quad \bar{\rho}_s \geq \frac{c_1^2 (2r)^{d/2}}{|\mathbb{T}_r^d|} = c_1^2 (2r)^{-d/2} \geq 2c_1^* \epsilon,$$

where c_1^* is a constant depending only on d .

Now we set $t = t(\epsilon) = c_2 r(\epsilon)^2 / \log r(\epsilon)$ and apply Lemma 2.3 to obtain

$$|\rho_t - \bar{\rho}_t| \leq c_1^* \epsilon,$$

when $c_2 = c_2(d)$ is chosen sufficiently small. Applying (16), we deduce that $\rho_t \geq c_1^* \epsilon$ for all $p \in [1/4, 1/2)$. Note that ρ_s is decreasing in s since particles can only disappear (see [13, Lemma 2] for a formal proof). Hence we have $\rho_s \geq c_1^* \epsilon$ for all $s \leq t$ and $p \in [1/4, 1/2)$. Therefore, for all $p \in [1/4, 1/2)$,

$$\mathbb{E}_p V_\infty \geq \mathbb{E}_p V_t = \int_0^t \rho_s ds \geq c_1^* \epsilon t \geq C \frac{\epsilon^{-4/d+1}}{-\log \epsilon}$$

for some constant $C > 0$ that may only depend on d . □

4.3. An upper bound on $\mathbb{E} V_t$ for DLAS on $\mathbb{Z}$. In this subsection, we prove Theorem 1.2, showing that occupation time of the origin in DLAS on $\mathbb{Z}$ is $O(t^{3/4})$ when $p = 1/2$ and $\lambda_B = 0$.

We will work with the sequential process defined in Section 2. Recall that in this process, we order the graph's vertices and then run the A -particles one at a time in sequence. Each A -particle runs until annihilation or time t . By Lemma 2.4, the total occupation time of the origin in this process is stochastically larger than the total occupation time of the origin in DLAS up to time t . Until the final proof of Theorem 1.2, we will work with the sequential process with particles present initially only on the positive integers, with vertices ordered $1, 2, \dots$. We call this the *one-sided sequential process*.

We start by proving an estimate on the probability of an A -particle at k reaching the origin. We then use this to bound the expected occupation time of the origin in the one-sided sequential process.

LEMMA 4.2. *Assume $p = 1/2$. Let G_k be the event that there is an A -particle initially at k and that it visits the root in the one-sided sequential process run for time t . Then for $1 \leq k \leq 2t$,*

$$\mathbb{P}(G_k) \leq C k^{-1/2} e^{-k^2/12t}$$

for an absolute constant C .

PROOF. We start by defining two random variables that are functions of the initial particle configuration. Fix k , and let $D = D[1, k-1]$ be the number of A -particles minus the number of B -particles initially in $[1, k-1]$. Observe that D is distributed as $2 \text{Bin}(k-1, 1/2) - (k-1)$. Consider the i th B -particle to the right of position k , and let L_i be its distance from k . Since L_i is a sum of i independent random variables with the geometric distribution of parameter $1/2$ on $\{1, 2, \dots\}$,

$$(17) \quad \mathbb{E} L_i = 2i.$$

Let $\mathcal{F}_{k-1}$ be the σ -algebra generated by the locations of all B -particles and the paths of A -particles starting at positions $1, \dots, k-1$ in the one-sided sequential process. This

represents the information available after running the process for the A -particles at positions $1, \dots, k-1$. Let F_k be the event that no B -particles remain at these sites and that there is an A -particle at position k . The event G_k cannot occur unless F_k occurs, since an A -particle at site k cannot move to the origin without colliding with one of these B -particles. If F_k occurs, then the key quantity is the distance to the right of k that the first remaining B -particle is found. Call this value R , setting it to 0 if F_k does not occur. Given R and that F_k occurs, the question is whether a simple random walk will reach the origin from k in t steps without moving to site $R+k$. Hence, using the gambler's ruin probability of reaching the origin without moving to site $R+k$ in any number of steps, we obtain the bound

$$\mathbb{P}(G_k \mid \mathcal{F}_{k-1}) \leq \frac{R}{R+k} \leq \frac{R}{k}$$

for all $k \geq 1$. When $3\sqrt{t} \leq k \leq 2t$, we can apply Proposition A.4 to obtain

$$\mathbb{P}(G_k \mid \mathcal{F}_{k-1}) \leq \frac{2R}{R+k} e^{-k^2/12t} \leq \frac{2R}{k} e^{-k^2/12t}.$$

Since $e^{-k^2/12t}$ is bounded away from 0 for $1 \leq k \leq 3\sqrt{t}$, it holds for all $1 \leq k \leq 2t$ that

$$(18) \quad \mathbb{P}(G_k \mid \mathcal{F}_{k-1}) = O\left(\frac{R}{k} e^{-k^2/12t}\right).$$

We claim that $R \leq L_{1+D}$ on the event F_k . Indeed, for F_k to occur, the A -particles initially in $[1, k-1]$ must annihilate all B -particles on that interval (and in particular $D \geq 0$). The remaining D A -particles can then annihilate at most the first D B -particles to the right of position k , which makes R at most the distance from k to the $(1+D)$ th B -particle to its right. And if F_k does not occur, then $R = 0$. Hence $R \leq L_{1+|D|}$ regardless of whether F_k occurs. Thus it follows from (18) that

$$\mathbb{P}(G_k \mid \mathcal{F}_{k-1}) = O\left(\frac{L_{1+|D|}}{k} e^{-k^2/12t}\right).$$

Hence

$$\begin{aligned} \mathbb{P}(G_k \mid D) &= \mathbb{E}\left[\mathbb{P}(G_k \mid \mathcal{F}_{k-1}) \mid D\right] \leq \mathbb{E}\left[\frac{L_{1+|D|}}{k} \mid D\right] O(e^{-k^2/12t}) \\ &= \frac{2(1+|D|)}{k} O(e^{-k^2/12t}), \end{aligned}$$

where the final equality uses (17) together with the independence of D from $(L_i)_{i \geq 1}$. Taking expectations gives

$$\begin{aligned} \mathbb{P}(G_k) &\leq \frac{1}{k} (1 + \mathbb{E}|D|) O(e^{-k^2/12t}) \leq \frac{1}{k} (1 + \sqrt{\mathbb{E}D^2}) O(e^{-k^2/12t}) \\ &= O(k^{-1/2} e^{-k^2/12t}). \end{aligned}$$

This finishes the proof. $\square$

PROPOSITION 4.3. *Let U_t^+ be the total occupation time of the origin in the one-sided sequential process run for time t . Then $\mathbb{E}U_t^+ = O(t^{3/4})$.*

PROOF. Let X_k be the total time that an A -particle starting at position k occupies the origin in the one-sided sequential process run for time t . As in Lemma 4.2, let G_k be the event that there is an A -particle initially at position k that visits the origin. Then

$$\mathbb{E}X_k = \mathbb{E}[X_k \mid G_k] \mathbb{P}(G_k) \leq t^{1/2} \mathbb{P}(G_k),$$

bounding $\mathbb{E}[X_k \mid G_k]$ by the occupation time of the origin up to time t by a random walk starting at the origin, which is at most $t^{1/2}$ by Lemma A.3.

We break $\mathbb{E}U_t^+$ into two parts:

$$(19) \quad \mathbb{E}U_t^+ = \sum_{k=1}^{\lfloor 2t \rfloor} \mathbb{E}X_k + \sum_{k=\lfloor 2t \rfloor+1}^{\infty} \mathbb{E}X_k \leq t^{1/2} \left(\sum_{k=1}^{\lfloor 2t \rfloor} \mathbb{P}(G_k) + \sum_{k=\lfloor 2t \rfloor+1}^{\infty} \mathbb{P}(G_k) \right).$$

For the first sum, we apply Lemma 4.2 and then bound the sum by an integral to get

$$(20) \quad \sum_{k=1}^{\lfloor 2t \rfloor} \mathbb{P}(G_k) \leq C \int_0^{\infty} x^{-1/2} e^{-x^2/12t} dx = C \int_0^{\infty} t^{1/4} u^{-1/2} e^{-u^2/12} du \\ = O(t^{1/4}).$$

For the second sum, we bound $\mathbb{P}(G_{\lfloor 2t \rfloor+1+i})$ by the probability of a random walk having displacement $2t+i$ in time t and apply Lemma A.2 to get

$$\mathbb{P}(G_{\lfloor 2t \rfloor+1+i}) \leq \exp\left(-\frac{(t+i)^2}{2(2t+i)}\right).$$

Hence

$$(21) \quad \sum_{k=\lfloor 2t \rfloor+1}^{\infty} \mathbb{P}(G_k) \leq \sum_{i=0}^{\infty} \exp\left(-\frac{(t+i)^2}{2(2t+i)}\right) \\ \leq \sum_{i=0}^{\infty} \exp\left(-\frac{(t+i)}{4}\right) \leq \int_{t-1}^{\infty} \exp(-x/4) dx = O(e^{-t/4}).$$

Equations (20) and (21) together with (19) prove the proposition. $\square$

REMARK 4.4. *It is possible to avoid the work of proving Proposition A.4 in the appendix as follows. First, use the usual gambler's ruin computation rather than Proposition A.4 in Lemma 4.2, proving only that $\mathbb{P}(G_k) = O(k^{-1/2})$. Then in (19), break $\mathbb{E}U_t^+$ into three sums, bounding the first using the estimate $\mathbb{P}(G_k) = O(k^{-1/2})$, the second using the moderate deviations estimate Lemma A.1 for the random walk, and the last using Lemma A.2 as was done. The downside of this approach is that it adds an extra logarithmic factor to the bound given in Proposition 4.3.*

Now we have all of the necessary estimates to prove our theorem.

PROOF OF THEOREM 1.2. Let V_t^+ and V_t^- denote the total occupation time of the origin by A -particles in DLAS on $\mathbb{Z}$ with $p = 1/2$ up to time t with particles placed initially only at positive integers and only at negative integers, respectively. Let V_t^0 denote the occupation time of the origin by A -particles in DLAS with only a single particle started at the origin (i.e., the local time of the origin by a single random walk if an A -particle is placed at the origin, and zero otherwise). Applying Lemma 2.5 twice, we have

$$\mathbb{E}V_t \leq \mathbb{E}V_t^+ + \mathbb{E}V_t^0 + \mathbb{E}V_t^-.$$

By Lemma 2.4 and Proposition 4.3, we have $\mathbb{E}V_t^+ \leq \mathbb{E}U_t^+ = O(t^{3/4})$, and by symmetry the same bound holds for $\mathbb{E}V_t^-$. Finally $\mathbb{E}V_t^0 \leq t^{1/2}$ by Lemma A.3, completing the proof that $\mathbb{E}V_t = O(t^{3/4})$. $\square$

4.4. *An upper bound on the critical exponent for DLAS on $\mathbb{Z}$.* In this section, we prove Theorem 1.4. As in our proof of Theorem 1.2, we will use the sequential version of DLAS defined in Section 2. Consider the sequential process run for infinite time with all particles in the initial configuration removed from the negative integers. Order the vertices $0, 1, 2, \dots$. Let U_k be 0 if a B -particle starts at site k ; otherwise, let it be the total time spent at the root by the A -particle initially at position k . If $k = 0$ and there is an A -particle at 0, we let U_0 be the number of visits to 0 by that particle for $t \geq 1$. The following estimate is most of the work in proving Theorem 1.4.

LEMMA 4.5. *For $k \geq 1$, let $D = D[0, k - 1]$. Then*

$$\mathbb{E}U_0 = \mathbf{1}\{0 \text{ contains an } A\text{-particle}\} + 2p(1 - p)^{-1} \mathbf{1}\{1 \text{ contains an } A\text{-particle}\}$$

and

$$\mathbb{E}U_k = 2p(1 - p)^{-1} \mathbb{E}[\mathbf{1}\{D \geq 0\}(1 + D)].$$

PROOF. We first describe the case $k \geq 1$. As in the proof of Lemma 4.2, let L_i be the distance from position k to the i th B -particle to its right. Let $L = \mathbf{1}\{D \geq 0\}L_{D+1}$. Let $\{k \rightarrow \mathbf{0}\}$ denote the event that there is an A -particle at k and it reaches the origin for some $t \geq 1$. In the sequential process we consider, each A -particle in $[0, k - 1]$ is eventually annihilated, since there are infinitely many B -particles in the initial configuration. Thus, if $D \geq 0$ and there is initially an A -particle at k , then the A -particles initially present in $[0, k - 1]$ will annihilate all B -particles in that region as well as the first D B -particles to the right of k . Hence, when the A -particle at k is run, there are no B -particles to its left, and the closest B -particle to its right is at distance L_{D+1} . By the gambler's ruin calculation,

$$\mathbb{P}(k \rightarrow \mathbf{0} \mid L) = \frac{pL}{k + L}.$$

Conditional on L , each time the particle visits the origin, it is annihilated without returning with probability

$$r = \frac{1}{2} \frac{1}{k + L}.$$

This is the probability it moves to the right on its first step and then reaches L before returning to the origin. Hence, conditional on $\{k \rightarrow \mathbf{0}\}$, the number of visits to $\mathbf{0}$ by the particle at k is distributed geometrically on the positive integers with success probability r . Since the particle stays at $\mathbf{0}$ for expected time 1 on each visit,

$$(22) \quad \mathbb{E}[U_k \mid L] = \frac{\mathbb{P}(k \rightarrow \mathbf{0} \mid L)}{r} = 2pL.$$

As in (17), we have $\mathbb{E}L_i = i/(1 - p)$, since L_i is a sum of i independent geometric random variables with success probability $1 - p$. Since $L = \mathbf{1}\{D \geq 0\}L_{1+D}$,

$$(23) \quad \mathbb{E}[L \mid D] = (1 - p)^{-1} \mathbf{1}\{D \geq 0\}(1 + D).$$

Taking expectations in (22) and (23) completes the proof for $k \geq 1$.

When $k = 0$ and there is an A -particle at 0, the situation is slightly different. The reason is that the particle will move to the negative integers and then eventually back to 0 a Geometric distributed with mean 1 number of times before moving to 1 for the first time. After doing so, if 1 contains a B -particle, then the A -particle is destroyed. If 1 contains an A -particle, then we may repeat the argument for U_1 with $D = 0$. This gives the claimed formula. $\square$

PROOF OF THEOREM 1.4. Let U^+ be the occupation time of $\mathbf{0}$ in the sequential process on the halfline run for infinite time, and observe that $U^+ = \sum_{k=1}^{\infty} U_k$. It suffices to show that $\mathbb{E}_p U^+ \leq C(1-2p)^{-3}$ for an absolute constant C , since then Lemmas 2.4 and 2.5 complete the proof as in the conclusion of the proof of Theorem 1.2.

Let $(S_k)_{k=0}^{\infty}$ be a simple random walk on the integers jumping a step in the positive direction with probability p and in the negative with probability $1-p$. By Lemma 4.5,

$$\mathbb{E}U^+ = \sum_{k=1}^{\infty} \mathbb{E}U_k = \frac{2p}{1-p} \sum_{k=1}^{\infty} \mathbb{E}[\mathbf{1}\{S_k \geq 0\}(1+S_k)].$$

This sum is essentially the expected area under the positive excursions of a random walk with negative drift, which we can compute exactly. We rewrite the sum to get

$$\mathbb{E}U^+ = \frac{2p}{1-p} \sum_{u=0}^{\infty} \mathbb{E} \left[\sum_{k=1}^{\infty} \mathbf{1}\{S_k \geq u\} \right] = \frac{2p}{1-p} \left(\sum_{u=0}^{\infty} \mathbb{E} \left[\sum_{k=0}^{\infty} \mathbf{1}\{S_k \geq u\} \right] - 1 \right).$$

Given that (S_k) ever hits u , the distribution of $\sum_{k=0}^{\infty} \mathbf{1}\{S_k \geq u\}$ is the same as the unconditional distribution of $\sum_{k=0}^{\infty} \mathbf{1}\{S_k \geq 0\}$. Hence, with M denoting the maximum value taken by (S_k) ,

$$(24) \quad \mathbb{E}U^+ = \frac{2p}{1-p} \left(\sum_{u=0}^{\infty} \mathbb{P}(M \geq u) \sum_{k=0}^{\infty} \mathbb{P}(S_k \geq 0) - 1 \right).$$

Now we compute the two sums. By the gambler's ruin calculation for biased random walk,

$$(25) \quad \sum_{u=0}^{\infty} \mathbb{P}(M \geq u) = \sum_{u=0}^{\infty} \left(\frac{p}{1-p} \right)^u = \frac{1-p}{1-2p}.$$

To approach the second sum, let S denote the number of steps for a simple random walk on the integers with bias p starting at 1 to hit 0. Since S is equal to 1 with probability $1-p$ and otherwise is distributed as the sum of two independent copies of itself, it satisfies $\mathbb{E}S = 1-p+2p\mathbb{E}S$ and hence $\mathbb{E}S = (1-p)/(1-2p)$. Now, let $T = \sum_{k=0}^{\infty} \mathbf{1}\{S_k \geq 0\}$. We claim that

$$(26) \quad T \stackrel{d}{=} \begin{cases} 1+S+T & \text{with probability } p, \\ 1+T & \text{with probability } p, \\ 1 & \text{with probability } 1-2p. \end{cases}$$

The first case corresponds to the event that (S_k) initially jumps to the right, which occurs with probability p . Then $1+S$ is contributed to the sum before it returns to 0, and the sum from that time on is distributed as T by the strong Markov property. In the second case, the first step of (S_k) is to the left (with probability $1-p$), and conditional on that event, it eventually returns to 0 (with probability $p/(1-p)$ by Gambler's ruin) and the sum for T resets. Hence T equals to $1+T$ with probability p . And in the final case, (S_k) never returns to 0. Taking expectations of both sides of (26) and solving the resulting equation yields

$$(27) \quad \mathbb{E}T = \frac{p(3-5p)}{(1-2p)^2} + 1.$$

From (24), (25), and (27),

$$\mathbb{E}U^+ = \frac{2p}{1-p} \left(\left(\frac{1-p}{1-2p} \right) \left(\frac{p(3-5p)}{(1-2p)^2} + 1 \right) - 1 \right) = \frac{2p^2(2-3p)^2}{(1-p)(1-2p)^3} = O((1-2p)^{-3}).$$

□

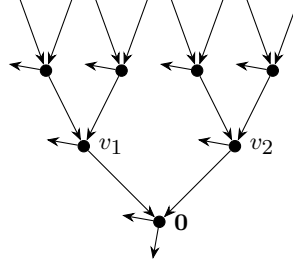


FIG 3. The tree $\vec{T}_4$, a 4-regular tree with oriented edges. Pictured is the portion of the tree that leads to the root 0 , which forms a binary tree.

5. Mean field behavior on bidirected trees. Consider the infinite $2d$ -regular tree with some vertex 0 distinguished as the root. At each vertex v , orient d of the edges to point toward v and d to point away. Define the random walk kernel K to send a walker at v along each of the d out-edges with probability $1/d$. We denote this oriented tree by $\vec{T}_{2d}$. In this section, we consider the two-type DLAS on $\vec{T}_{2d}$ with $\lambda_B = 0$. As usual, the initial configuration is one particle per site, where each particle is independently given type A with probability p . The subset of $\vec{T}_{2d}$ made up of vertices with a path to the root forms a rooted d -ary tree (see Figure 3). We will ignore the rest of the tree, since no particles from it can contribute to V_t . We define level k of the tree as the d^k vertices in this subtree from which there is a directed path of k edges toward the root.

By Lemma 2.2, when bounding $\mathbb{E}V_t$ we can ignore particles far from the root. Thus, our strategy will be to work with DLAS with particles removed beyond some level n (which we will later take to be $\lceil ct \rceil$) and to ignore time, counting the total number of visits to the root by A -particles in any amount of time. The main goal of this section is to prove the following bounds on this quantity.

PROPOSITION 5.1. *Consider DLAS on $\vec{T}_{2d}$ with particles initially placed only at levels $1, \dots, n$. Let W_n denote the number of A -particles that visit the root in any amount of time.*

(a) *If $p = 1/2$, then for absolute constants c and C , it holds for all $d, n \geq 2$ that*

$$c \log n \leq \mathbb{E}W_n \leq C \log n.$$

(b) *If $\epsilon = 1/2 - p > 0$, then $\mathbb{E}W_n$ approaches a finite limit as $n \rightarrow \infty$, and for absolute constants c, C , and η , it holds for all $0 < \epsilon < \eta$ and all $d \geq 2$ that*

$$c \log\left(\frac{1}{\epsilon}\right) \leq \lim_{n \rightarrow \infty} \mathbb{E}W_n \leq C \log\left(\frac{1}{\epsilon}\right).$$

Before we go further, we show how these bounds prove Theorems 1.5 and 1.6. First, we invoke Lemmas 2.1 and 2.2 to relate V_t and W_n .

LEMMA 5.2. *For some absolute constant c ,*

$$\mathbb{E}W_{\lfloor t/2 \rfloor} - O(1) \leq \mathbb{E}V_t \leq \mathbb{E}W_{\lceil ct \rceil} + O(1).$$

PROOF. We start with the upper bound. Let c be the constant from Lemma 2.2. Fix $t \geq 0$, and for each $s \geq 0$, let $\bar{\rho}_s$ and $\bar{V}_s$ respectively denote the density and the total occupation time

at time s of the root by A -particles in DLAS on $\tilde{\mathcal{T}}_{2d}$ with particles beyond level ct removed from the initial configuration. Then by Lemma 2.2,

$$\mathbb{E}V_t \leq \mathbb{E}\bar{V}_t + \int_0^t |\rho_s - \bar{\rho}_s| ds \leq \mathbb{E}\bar{V}_\infty + te^{-t}.$$

Since any A -particle visiting the root is either immediately annihilated by a B -particle or remains there for expected time 1, we have $\mathbb{E}\bar{V}_\infty \leq \mathbb{E}W_{[ct]} + p$ (the reason for the extra p is that $\bar{V}_\infty$ includes the possible contribution that an A -particle initially at the root provides to the root's occupation time.) This completes the proof of the upper bound.

For the lower bound, let $\tilde{V}_t$ denote the total occupation time of the root by A -particles in DLAS on $\tilde{\mathcal{T}}_{2d}$ with the starting configuration containing A -particles only at levels $1, \dots, \lfloor t/2 \rfloor$. By Lemma 2.1, we have $\mathbb{E}V_t \geq \mathbb{E}\tilde{V}_t$. Let $\tilde{N}_t$ be the total number of A -particles in this truncated system that visit the root in time t . Since each A -particle moving to the root stays there for expected time 1 except possibly the first one (which is immediately annihilated if the root initially contains a B -particle), we have $\mathbb{E}\tilde{V}_t \geq \mathbb{E}\tilde{N}_t - 1$. Thus it suffices to show that $\mathbb{E}\tilde{N}_t \geq \mathbb{E}W_{\lfloor t/2 \rfloor} - O(1)$.

Let $\tilde{N}$ be the number of A -particles in this truncated initial configuration whose underlying random walk paths visit the root, whether or not they are annihilated before reaching it. There are d^k vertices at level k , and a random walk starting at such a vertex has probability $1/d$ of moving toward the root at each step and hence has probability d^{-k} of ever visiting the root. By expressing $\tilde{N}$ as a sum of indicators, we find that $\mathbb{E}\tilde{N} = O(t)$ and $\mathbb{E}\tilde{N}^2 = O(t^2)$. Now, let E be the event that the underlying random walk for each particle counted by $\tilde{N}$ makes at least $t/2$ jumps in time t . On this event, we have $\tilde{V}_t = W_{\lfloor t/2 \rfloor}$, since all particles that ever visit the origin will do so in time t . It follows that

$$\tilde{V}_t \geq W_{\lfloor t/2 \rfloor} \mathbf{1}_E \geq W_{\lfloor t/2 \rfloor} - \tilde{N} \mathbf{1}_{E^c}.$$

Now we bound $\mathbb{E}[\tilde{N} \mathbf{1}_{E^c}]$. The probability that a random walk path jumps fewer than $t/2$ times in time t is bounded by e^{-bt} for some $b > 0$, by basic concentration properties of the Poisson distribution. Thus $\mathbb{P}(E^c \mid \tilde{N}) \leq \tilde{N}e^{-bt}$. This yields $\mathbb{E}[\tilde{N} \mathbf{1}_{E^c}] \leq e^{-bt} \mathbb{E}\tilde{N}^2 = O(1)$, and therefore $\mathbb{E}\tilde{V}_t \geq \mathbb{E}W_{\lfloor t/2 \rfloor} - O(1)$. $\square$

PROOF OF THEOREM 1.5. The theorem follows immediately by applying Lemma 5.2 followed by Proposition 5.1(a). $\square$

PROOF OF THEOREM 1.6. Let $\epsilon = 1 - 2p$. By Lemma 5.2 and Proposition 5.1(b),

$$\limsup_{t \rightarrow \infty} \mathbb{E}_p V_t \leq C \log(1/\epsilon) + O(1) \leq C' \log(1/\epsilon)$$

and

$$\liminf_{t \rightarrow \infty} \mathbb{E}_p V_t \geq c \log(1/\epsilon) - O(1) \geq c' \log(1/\epsilon)$$

for sufficiently small ϵ . Since V_t converges upwards to V_∞ , by the monotone convergence theorem $\mathbb{E}_p V_\infty$ exists and is bounded above by $C' \log(1/\epsilon)$ and below by $c' \log(1/\epsilon)$. $\square$

5.1. A recursive distributional equation for W_n . We now set out to prove Proposition 5.1. The main idea to understand W_n is to take advantage of the recursive structure of the tree. Let $v_1, \dots, v_d$ be the vertices at level 1 of the tree (see Figure 3). For any of these vertices, the number of A -particles visiting them in this system is distributed as W_{n-1} . Because the tree is directed, these counts are independent. Thus, we obtain a distributional equation expressing the distribution of W_n in terms of d independent copies of W_{n-1} as

follows. Let $W_{n-1}^{(1)}, \dots, W_{n-1}^{(d)}$ denote these counts. Let Y_i be 1 if v_i initially contains an A -particle and be -1 if it initially contains a B -particle. From the dynamics of the system, the number of A -particles with a chance of moving from v_i to the root is $(W_{n-1}^{(i)} + Y_i)^+$, where $(\cdot)^+ = \max\{0, \cdot\}$. Each of these A -particles independently moves to the root with probability d^{-1} . Hence,

$$(28) \quad W_n \stackrel{d}{=} \sum_{i=1}^d \text{Bin}\left((W_{n-1}^{(i)} + Y_i)^+, d^{-1}\right).$$

Here we use the notation $\text{Bin}(X, p)$ to denote the p -thinning of a nonnegative integer-valued random variable X . That is, $\text{Bin}(X, p)$ is defined as $\sum_{i=1}^X B_i$, where $B_1, B_2, \dots$ are independent of each other and X and have distribution $\text{Ber}(p)$.

Notationally, we will express this relationship between W_{n-1} and W_n by defining an operator $\mathcal{A}$ on probability distributions that maps the law of W_{n-1} to the law of W_n . Let $Y_1, \dots, Y_d$ be independent, each equal to 1 with probability p and -1 with probability $1 - p$. Let $X_1, \dots, X_d$ be independent copies of an arbitrary random variable X taking values in the nonnegative integers. Then we define the result of applying $\mathcal{A}$ to the law of X by

$$(29) \quad \mathcal{A}X = \sum_{i=1}^d \text{Bin}\left((X_i + Y_i)^+, d^{-1}\right).$$

We will sometimes abuse notation and treat $\mathcal{A}X$ as a random variable with this distribution. The following lemma summarizes the fact given in (28):

LEMMA 5.3. *The random variables $(W_n)_{n \geq 0}$ from Proposition 5.1 satisfy the distributional equality*

$$W_{n+1} \stackrel{d}{=} \mathcal{A}W_n.$$

This operator is monotone with respect to the stochastic order.

LEMMA 5.4. *If $X \preceq X'$, then $\mathcal{A}X \preceq \mathcal{A}X'$.*

PROOF. Let

$$\mathcal{A}X = \sum_{i=1}^d \text{Bin}\left((X_i + Y_i)^+, d^{-1}\right)$$

and

$$\mathcal{A}X' = \sum_{i=1}^d \text{Bin}\left((X'_i + Y_i)^+, d^{-1}\right),$$

with $X_1, \dots, X_d$ i.i.d. copies of X and $X'_1, \dots, X'_d$ i.i.d. copies of X' . Since $X \preceq X'$, we can couple each X_i with X'_i so that $X_i \leq X'_i$ a.s., thus producing a coupling in which $\mathcal{A}X \leq \mathcal{A}X'$ a.s. $\square$

Our goal now is to use the operator $\mathcal{A}$ to analyze the growth of $\mathbb{E}W_n$. We make the following observation:

LEMMA 5.5. *Let X be a random variable taking nonnegative integer values. Then*

$$(30) \quad \mathbb{E}[\mathcal{A}X] = \mathbb{E}X + 2p - 1 + (1 - p)\mathbb{P}(X = 0).$$

PROOF. This follows from the definition of $\mathcal{A}$ and the statement

$$\mathbb{E}[(X_i + Y_i)^+] = \mathbb{E}[X_i + Y_i + \mathbf{1}\{X_i = 0, Y_i = -1\}] = \mathbb{E}X + \mathbb{E}Y_i + \mathbb{P}(X = 0)\mathbb{P}(Y_i = -1),$$

where $X_1, \dots, X_d$ and $Y_1, \dots, Y_d$ are independent with $X_i \stackrel{d}{=} X$ and Y_i equal to 1 with probability p and to -1 with probability $1 - p$. $\square$

This shows that the growth of W_n depends on its concentration, in that $\mathbb{P}(W_n = 0)$ is the key quantity in (30) when we apply it to bound $\mathbb{E}W_{n+1} - \mathbb{E}W_n$. Heuristically, the explanation for the growth rate of $\mathbb{E}W_n$ is that W_n has a Poisson-like lower tail, and in particular $\mathbb{P}(W_n = 0) \approx e^{-c\mathbb{E}W_n}$. In the critical case $p = 1/2$, Lemma 5.5 then gives

$$\mathbb{E}W_{n+1} - \mathbb{E}W_n \approx Ce^{-c\mathbb{E}W_n}.$$

Solving this difference equation shows that $\mathbb{E}W_n$ grows logarithmically. In the subcritical case, a similar difference equation shows that $\mathbb{E}W_n$ converges as n tends to infinity and gives bounds on the limit.

The lower bounds in Proposition 5.1 are proven just as in this description. We need an anticoncentration bound of the form $\mathbb{P}(W_n = 0) \geq Ce^{-c\mathbb{E}W_n}$. To get this, we consider the worst-case scenario, when W_n is as concentrated as possible. One can then calculate that even in this case, $\mathcal{A}W_n$ is not too concentrated, and therefore W_{n+1} satisfies the desired anticoncentration bound.

The upper bounds in Proposition 5.1 are more difficult. For technical reasons, we bound the expectations of a sequence of random variables $(U_n)_{n \geq 0}$ that serves as a stochastic upper bound for $(W_n)_{n \geq 0}$. We then show an exponential concentration bound for the lower tail of U_n using the theory of size-bias couplings.

In Sections 5.2 and 5.3, we lay out some background material we will need for this argument. The upper bounds from Proposition 5.1 are proven in Section 5.4 followed by the lower bounds in Section 5.5.

5.2. Concentration by size-bias coupling. Let π be a probability distribution on the non-negative real numbers with mean μ . The *size-bias transform* of π is the distribution π^s with Radon–Nikodym derivative

$$\frac{d\pi^s(x)}{d\pi} = \frac{x}{\mu}.$$

We will often abuse notation and speak of the size-bias transform of a random variable as another random variable, rather than referring to their distributions. In this spirit, the size-bias transform of a random variable X on the nonnegative integers is the random variable X^s with distribution given by

$$\mathbb{P}(X^s = k) = \frac{k}{\mathbb{E}X} \mathbb{P}(X = k).$$

We will in general use the notation X^s to denote the size-bias transform of X .

The size-bias transform comes up in a variety of contexts; see [4] for a broad survey. Our interest here has its roots in Stein’s method for distributional approximation. If a random variable and its size-bias transform are close to each other in the right sense, one can use Stein’s method to prove that the random variable is approximately Gaussian or Poisson (see [38, Sections 3.4 and 4.3]). It was eventually realized that the same approach could be used to prove concentration inequalities [22, 3, 16]. The starting point is that if X and its size-bias transform exactly satisfy $X^s \stackrel{d}{=} X + 1$, then X is Poisson [4, Corollary 11.3]. If instead $X^s \preceq X + 1$, then it can be proven that X satisfies a Poisson-like tail concentration bound

[3, Theorems 1.2–1.3]. We use a result that relaxes the condition $X^s \preceq X + 1$ further. A coupling of X and its size-bias transform X^s is (c, p) -bounded for the upper tail if

$$\mathbb{P}(X^s \leq X + c \mid X^s) \geq p \text{ a.s.},$$

and is (c, p) -bounded for the lower tail if

$$\mathbb{P}(X^s \leq X + c \mid X) \geq p \text{ a.s.}$$

In this language, the condition $X^s \preceq X + 1$ is that X admits a size-bias coupling that is $(1, 1)$ -bounded for the upper and lower tails.

PROPOSITION 5.6 ([16, Theorem 3.3]). *Let $\mu = \mathbb{E}X$ and let $h(x) = (1+x)\log(1+x) - x$ for $x \geq -1$.*

(a) *If X admits a size-bias coupling that is (c, p) -bounded for the upper tail, then for all $x \geq 0$,*

$$\mathbb{P}(X - \mu/p \geq x) \leq \exp\left(-\frac{\mu}{cp}h\left(\frac{px}{\mu}\right)\right) \leq \exp\left(-\frac{x^2}{2c(x/3 + \mu/p)}\right).$$

(b) *If X admits a size-bias coupling that is (c, p) -bounded for the lower tail, then for all $0 \leq x \leq p\mu$,*

$$\mathbb{P}(X - p\mu \leq -x) \leq \exp\left(-\frac{p\mu}{c}h\left(-\frac{x}{p\mu}\right)\right) \leq \exp\left(-\frac{x^2}{2pc\mu}\right).$$

In particular, under the hypotheses of part (b) of this theorem, we have $\mathbb{P}(X = 0) \leq \exp(-p\mu/2c)$ by setting $x = p\mu$. (In [16], part (b) of this theorem is stated with the condition $x < p\mu$, but taking a limit extends it to $x = p\mu$.)

To use this theorem, we will need to compute size-bias transforms. In general, this can be done effectively for random variables that are sums of simpler random variables (e.g., indicators), which need not be independent (see [4, Section 2.4]). The case of a size-bias transform of an independent sum is the simplest one: just choose a random summand with probability in proportion to its expectation, and then apply the transform to that summand only. The following lemma is a formal statement of this:

LEMMA 5.7 ([4, eq. (25)]). *Let $S = X_1 + \cdots + X_n$, where $X_1, \dots, X_n$ are nonnegative independent random variables. Choose I from $\{1, \dots, n\}$ independently of all else, choosing $I = i$ with probability $\mathbb{E}X_i/\mathbb{E}S$. Then*

$$S^s \stackrel{d}{=} X_1 + \cdots + X_n + X_I^s - X_I.$$

Next, we compute the size-bias transform of a p -thinning:

LEMMA 5.8. *Let X be a nonnegative random variable with finite expectation. Let $Y = \text{Bin}(X, p)$. Then*

$$(31) \quad Y^s \stackrel{d}{=} 1 + \text{Bin}(X^s - 1, p).$$

PROOF. Rewrote entire proof with more details –TJ To put this statement in a more general context, suppose Y is a mixture of random variables Z_n governed by a probability measure h ; that is, for Borel sets $\mathcal{B}$,

$$\mathbb{P}(Y \in \mathcal{B}) = \int \mathbb{P}(Z_n \in \mathcal{B}) h(dn).$$

Then Y^s is the mixture of the transforms Z_n^s , but governed not by h but by h^s , the measure whose Radon–Nikodym derivative satisfies $\frac{dh^s}{dh}(n) = \frac{\mathbb{E}Z_n}{\mathbb{E}Y}$ [4, Lemma 2.4].

In our case, Y is the mixture of $Z_n \sim \text{Bin}(n, p)$ governed by the law of X . By Lemma 5.7, the size-bias transform of Z_n is $1 + \text{Bin}(n-1, p)$. Since $\mathbb{E}Z_n$ is proportional to n , the measure h' is the size-bias transform of the law of X . Thus Y^s is the mixture of $1 + \text{Bin}(n-1, p)$ governed by the law of X^s , as stated in (31). $\square$

Finally, size-biasing ignores any mass the distribution places on 0:

LEMMA 5.9. *Let $X^>$ denote a random variable distributed as X conditioned on $X > 0$. Then $(X^>)^s \stackrel{d}{=} X^s$.*

PROOF. This follows immediately from the definition of the size-bias transform. $\square$

The size-bias transform does not mesh well with the usual stochastic order, in that it is not true in general that $X \preceq Y$ implies that $X^s \preceq Y^s$. But this is true in a stronger stochastic order known as the *likelihood ratio order*. For integer-valued random variables X and Y , we say that $X \preceq_{lr} Y$ if $\mathbb{P}(Y = k)/\mathbb{P}(X = k)$ is increasing over the union of the supports of X and Y , interpreting this quantity as ∞ when $\mathbb{P}(Y = k) > 0$ and $\mathbb{P}(X = k) = 0$. It is not hard to show that $X \preceq_{lr} Y$ implies $X \preceq Y$ [39, Theorem 1.C.1].

PROPOSITION 5.10. *Let X and Y take values in the nonnegative integers. If $X \preceq_{lr} Y$, then $X^s \preceq_{lr} Y^s$.*

PROOF. By definition of the size-bias transform,

$$\frac{\mathbb{P}(Y^s = n)}{\mathbb{P}(X^s = n)} = \frac{(\mathbb{E}X)\mathbb{P}(Y = n)}{(\mathbb{E}Y)\mathbb{P}(X = n)}.$$

This is increasing in n since $X \preceq_{lr} Y$. $\square$

5.3. *Log-concave random variables.* Let X be a random variable taking values in the nonnegative integers, and let $P_n = \mathbb{P}(X = n)$. We say that X and its distribution are *log-concave* if

- (i) $P_n^2 \geq P_{n-1}P_{n+1}$ for all $n \geq 1$; and
- (ii) the sequence $P_0, P_1, \dots$ has no internal zeros (i.e., if $P_i, P_k > 0$ for $i < k$, then $P_j > 0$ for all $i \leq j \leq k$).

Our need for log-concave random variables boils down to the following fact, which will be used in combination with Proposition 5.10 at a key moment.

LEMMA 5.11. *Let X be a random variable taking values in the nonnegative integers. Then $X \preceq_{lr} X + 1$ if and only if X is log-concave.*

PROOF. Let $P_n = \mathbb{P}(X = n)$ as above. If $(P_n)_{n \geq 1}$ has no internal zeros and N is the highest value such that $P_N > 0$ (allowing $N = \infty$), then the statements

$$P_{n-1}P_{n+1} \leq P_n^2 \quad \text{for all } n \geq 1$$

and

$$\frac{P_{n-1}}{P_n} \leq \frac{P_n}{P_{n+1}} \quad \text{for all } 1 \leq n \leq N$$

are equivalent. The first of these statements is the log-concavity of $(P_n)_{n \geq 0}$, and the second is the meaning of $X \preceq_{\text{lr}} X + 1$ under the assumption of no internal zeros. Thus, the log-concavity of X implies $X \preceq_{\text{lr}} X + 1$. For the other direction, we just need to show that $X \preceq_{\text{lr}} X + 1$ implies that $(P_n)_{n \geq 0}$ has no internal zeros. Indeed, suppose $P_i, P_j > 0$ for $i < j$. Since $P_i/P_{i+1} \leq P_{j-1}/P_j$, we have $P_{i+1} > 0$, and then we can proceed to show that $P_{i+2} > 0$, and so on. $\square$

Next, we state a few technical facts about log-concave random variables, which come from translating combinatorial results into probabilistic ones. The first result we give is well known; see [41, Proposition 2] for the most standard proof, or see [31] for a completely elementary proof.

PROPOSITION 5.12. *If X and Y are log-concave and independent, then so is $X + Y$.*

Next, we show that log-concavity is preserved under thinning, which follows quickly from a combinatorial result of Brenti's [11]. We suspect that this has been used before, but we could not find it stated anywhere.

PROPOSITION 5.13. *Let X be log-concave. Then $\text{Bin}(X, p)$ is also log-concave for any $0 < p < 1$.*

PROOF. First, we prove this under the assumption that X has finite support. Let $P_n = \mathbb{P}(X = n)$ as before. We call a sequence $x_0, x_1, \dots$ log-concave if it satisfies the same conditions as $P_0, P_1, \dots$, i.e., no internal zeros and $x_n^2 \geq x_{n-1}x_{n+1}$ for $n \geq 1$. From the log-concavity of $(P_n)_{n \geq 0}$, it follows by direct calculation that the sequence $\{(1-p)^n P_n\}_{n \geq 0}$ is also log-concave. Now, let

$$Q_k = \sum_{n=k}^{\infty} \binom{n}{k} (1-p)^n P_n.$$

By [11, Theorem 2.5.3], the sequence $(Q_k)_{k \geq 0}$ is log-concave (this is where we use that X has finite support). This implies that the sequence

$$\left\{ \left(\frac{p}{1-p} \right)^k Q_k \right\}_{k \geq 0}$$

is log-concave, by direct calculation. And now we have proven that $\text{Bin}(X, p)$ is log-concave, since

$$\mathbb{P}(\text{Bin}(X, p) = k) = \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \mathbb{P}(X = n) = \left(\frac{p}{1-p} \right)^k Q_k.$$

Finally, we remove the condition that X has finite support with a limit argument. Let X_n be distributed as X conditioned on $X \leq n$. Then $X_n \rightarrow X$ in distribution. Since X_n is log-concave and has finite support, $\text{Bin}(X_n, p)$ is log-concave. It is then straightforward to see that $\text{Bin}(X_n, p)$ converges in distribution to $\text{Bin}(X, p)$ and that a weak limit of log-concave random variables is log-concave. $\square$

5.4. *Upper bounds on root visits.* As we mentioned at the end of Section 5.1, we will define a sequence $(U_n)_{n \geq 0}$ that serves as an upper bound for $(W_n)_{n \geq 0}$. Let $U_0 = 1$ a.s. Then inductively define the sequence by

$$U_{n+1} \stackrel{d}{=} \mathcal{A}U_n^>,$$

where we use the notation $X^>$ as in Lemma 5.9. The overarching goal of the section is to show that U_n admits a size-bias coupling bounded for the lower tail, which we do by induction. As we described following Lemma 5.5, we can use a concentration bound for the lower tail of U_n to bound the growth of $\mathbb{E}U_n$, which will allow us to prove $\mathbb{E}U_n = O(\log n)$. Finally, the same estimate holds for W_n since it is bounded by U_n . We start by proving that U_n is in fact an upper bound for W_n .

LEMMA 5.14. *For all $n \geq 0$,*

$$W_n \preceq U_n.$$

PROOF. We prove this by induction, starting with $0 = W_0 \preceq U_0 = 1$. Now suppose that $W_n \preceq U_n$. Since $U_n \preceq U_n^>$, we have $W_n \preceq U_n^>$, and applying $\mathcal{A}$ to both sides of this inequality yields $W_{n+1} \preceq U_{n+1}$ by Lemmas 5.3 and 5.4. $\square$

The point of working with (U_n) rather (W_n) is that we will be able to show that the random variables (U_n) are log-concave. First we give a technical lemma in this direction.

LEMMA 5.15. *Let $Y_1, \dots, Y_d$ be i.i.d. random variables taking values ± 1 with $\mathbb{P}(Y_i = 1) = p$. For $d \geq 2$ and $p \geq 4/9$, the random variable*

$$(32) \quad \text{Bin}\left(\sum_{i=1}^d (Y_i + 1), d^{-1}\right)$$

is log-concave.

PROOF. First, consider the case $d \geq 3$. We claim that for $p \geq 3/7$, the distribution $\text{Bin}(Y_i + 1, 1/3)$ is log-concave. Indeed, let X have this distribution. We only need to check that $\mathbb{P}(X = 1)^2 \geq \mathbb{P}(X = 0)\mathbb{P}(X = 2)$. A computation shows that

$$\mathbb{P}(X = 1)^2 - \mathbb{P}(X = 0)\mathbb{P}(X = 2) = \frac{(7p - 3)p}{27},$$

confirming the claim. For $x \leq 1/3$, the distribution $\text{Bin}(Y_i + 1, x)$ is a thinning of $\text{Bin}(Y_i + 1, 1/3)$ and hence is log-concave by Proposition 5.13. By viewing (32) as

$$\sum_{i=1}^d \text{Bin}(Y_i + 1, d^{-1}),$$

we see that it is a sum of independent log-concave random variables for any $d \geq 3$ and by Proposition 5.12 is log-concave.

Now we turn to the case $d = 2$. Let X be distributed as (32), and let $P_k = \mathbb{P}(X = k)$ for $k = 0, \dots, 4$. We then compute directly

$$\begin{aligned} P_1^2 - P_0P_2 &= \frac{(9p - 4)(3p - 4)^2p}{128}, \\ P_2^2 - P_1P_3 &= \frac{\left(13\left(p - \frac{12}{13}\right)^2 + \frac{64}{13}\right)p^2}{64}, \\ P_3^2 - P_2P_4 &= \frac{(9p - 4)p^2}{128}, \end{aligned}$$

all of which are positive for $p \geq 4/9$. $\square$

PROPOSITION 5.16. *If $p \geq 4/9$, then U_n is log-concave for all $n \geq 0$.*

PROOF. We prove this by induction. Trivially, U_0 is log-concave. Now, assume that U_n is log-concave. This implies that $U_n^>$ is log-concave, since the sequence $(\mathbb{P}(U_n^> = k))_{k \geq 0}$ is a rescaled version of the sequence $(\mathbb{P}(U_n = k))_{k \geq 0}$ with the $k = 0$ term set to zero. Let $X_1, \dots, X_d$ be i.i.d. copies of $U_n^>$, and let $Y_1, \dots, Y_d$ be i.i.d. random variables taking values ± 1 with $\mathbb{P}(Y_i = 1) = p$. Looking back at the definition of $\mathcal{A}$ in (29), the distribution of U_{n+1} is given by

$$(33) \quad \mathcal{A}U_n^> = \text{Bin}\left(\sum_{i=1}^d (X_i + Y_i)^+, d^{-1}\right) = \text{Bin}\left(\sum_{i=1}^d (X_i + Y_i), d^{-1}\right)$$

with the replacement of $(X_i + Y_i)^+$ by $X_i + Y_i$ justified by the strict positivity of X_i . Now we take advantage of this positivity to break $X_i + Y_i$ into nonnegative terms $X_i - 1$ and $Y_i + 1$, giving us

$$(34) \quad \begin{aligned} \mathcal{A}U_n^> &= \text{Bin}\left(\sum_{i=1}^d (X_i - 1) + \sum_{i=1}^d (Y_i + 1), d^{-1}\right) \\ &= \text{Bin}\left(\sum_{i=1}^d (X_i - 1), d^{-1}\right) + \text{Bin}\left(\sum_{i=1}^d (Y_i + 1), d^{-1}\right). \end{aligned}$$

Since X_i is log-concave, so is $X_i - 1$. Thus the first term on the right-hand side in (34) is a thinned sum of log-concave random variables and hence is log-concave by Propositions 5.12 and 5.13. The second term is log-concave by Lemma 5.15. A final application of Proposition 5.12 shows that (34) is a log-concave distribution, completing the induction. $\square$

Now, we start setting up an induction to show that U_n admits a size-bias coupling that is $(1, q)$ -bounded for the lower tail for $q > 0$.

LEMMA 5.17. *Let $X_1, \dots, X_d$ be distributed as $U_n^>$, and let $Y_1, \dots, Y_d$ be i.i.d. random variables taking values ± 1 with $\mathbb{P}(Y_i = 1) = p \geq 4/9$, and let all be independent of each other. Let U_n^s be independent of $X_2, \dots, X_d$ as well. For all $n \geq 0$,*

$$U_{n+1}^s \preceq 1 + \text{Bin}\left(U_n^s + (Y_1 + 1)^s - 1, d^{-1}\right) + \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right)$$

PROOF. As in (33),

$$U_{n+1} \stackrel{d}{=} \sum_{i=1}^d \text{Bin}(X_i + Y_i, d^{-1}).$$

By Lemma 5.7, we obtain the size-bias transform of U_{n+1} by choosing a term in this sum at random and biasing it. Since all terms are identically distributed, we can just bias the first term. Applying Lemma 5.8 to this term, we obtain

$$(35) \quad U_{n+1} \stackrel{d}{=} 1 + \text{Bin}((X_1 + Y_1)^s - 1, d^{-1}) + \sum_{i=2}^d \text{Bin}(X_i + Y_i, d^{-1}).$$

Now, we stochastically bound $(X_1 + Y_1)^s$. We first rewrite it as $((X_1 - 1) + (Y_1 + 1))^s$, noting that $X_1 - 1$ and $Y_1 + 1$ are both nonnegative. By Lemma 5.7,

$$(36) \quad (X_1 + Y_1)^s \stackrel{d}{=} \begin{cases} (X_1 - 1)^s + (Y_1 + 1) & \text{with probability } a, \\ (X_1 - 1) + (Y_1 + 1)^s & \text{with probability } 1 - a, \end{cases}$$

for some a that we will avoid calculating.

Now, we show that both parts of this mixture are stochastically dominated by $U_n^s + (Y_1 + 1)^s$. First observe that U_n is log-concave by Proposition 5.16, and therefore so is $X_1 \stackrel{d}{=} U_n^>$. Thus $X_1 - 1$ is log-concave, and $X_1 - 1 \preceq_{\text{lr}} X_1$ by Lemma 5.11. Applying Proposition 5.10 and the fact that dominance in the lr-order implies dominance in the usual order, we see that $(X_1 - 1)^s \preceq X_1^s \stackrel{d}{=} U_n^{>s}$. Finally, $U_n^{>s} \stackrel{d}{=} U_n^s$ by Lemma 5.9, establishing that $(X_1 - 1)^s \preceq U_n^s$. Next, we have $Y_1 + 1 \preceq (Y_1 + 1)^s$, a fact about size-biasing (it holds because $X \preceq_{\text{lr}} X^s$ in general by direct calculation). This completes the proof that the first part of the mixture is dominated by $U_n^s + (Y_1 + 1)^s$. To show this for the second part, we simply note that $X_1 - 1 \preceq (X_1 - 1)^s \preceq U_n^s$.

Since $U_n^s + (Y_1 + 1)^s$ stochastically dominates both parts of the mixture in (36), it stochastically dominates $(X_1 + Y_1)^s$ as well. Applying this to (35) completes the proof. $\square$

Now we show the existence of our size-bias coupling for U_n . The idea is to iterate the stochastic relation proven in Lemma 5.17 to obtain a stochastic relation between U_n and U_n^s that implies the existence of the coupling.

PROPOSITION 5.18. *Let $p \geq 4/9$ and $q = \prod_{i=1}^{\infty} (1 - 2^{-i})^2 \approx .083$. For all $n \geq 0$, the random variable U_n admits a size-bias coupling that is $(1, q)$ -bounded for the lower tail.*

PROOF. The bulk of the proof is to show that for all $n \geq 1$,

$$(37) \quad U_n^s \preceq 1 + \sum_{i=1}^n \text{Bin}(2, d^{-i}) + U_n,$$

which we prove by induction. Let $X_1, \dots, X_d$ be distributed as $U_n^>$, and let $Y_1, \dots, Y_d$ be i.i.d. random variables taking values ± 1 with $\mathbb{P}(Y_i = 1) = p$, all independent of each other. For the base case, set $n = 1$. Then by Lemma 5.17,

$$\begin{aligned} U_1^s &\preceq 1 + \text{Bin}(U_0^s + (Y_1 + 1)^s - 1, d^{-1}) + \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right) \\ &\preceq 1 + \text{Bin}(2, d^{-1}) + \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right) \\ &\preceq 1 + \text{Bin}(2, d^{-1}) + \text{Bin}\left(\sum_{i=1}^d (X_i + Y_i), d^{-1}\right) \stackrel{d}{=} 1 + \text{Bin}(2, d^{-1}) + U_1. \end{aligned}$$

From the first to the second line, we have applied the inequality $U_0^s + (Y_1 + 1)^s - 1 \leq 2$, since $U_0^s = 1$ and $(Y_1 + 1)^s \leq 2$. The last equality is due to the definition of U_1 .

Now, we assume that (37) holds for n , and we show it holds for $n + 1$. Apply Lemma 5.17 together with the inductive hypothesis to obtain

$$U_{n+1}^s \preceq 1 + \text{Bin}\left(1 + \sum_{i=1}^n \text{Bin}(2, d^{-i}) + U_n + (Y_1 + 1)^s - 1, d^{-1}\right)$$

$$+ \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right).$$

Again, we have $(Y_1 + 1)^s \leq 2$. We can also apply the bound $U_n \preceq U_n^> \stackrel{d}{=} X_1$. This yields

$$\begin{aligned} U_{n+1}^s &\preceq 1 + \text{Bin}\left(\sum_{i=1}^n \text{Bin}(2, d^{-i}) + X_1 + 2, d^{-1}\right) + \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right) \\ &\stackrel{d}{=} 1 + \sum_{i=1}^{n+1} \text{Bin}(2, d^{-i}) + \text{Bin}(X_1, d^{-1}) + \text{Bin}\left(\sum_{i=2}^d (X_i + Y_i), d^{-1}\right) \\ &\preceq 1 + \sum_{i=1}^{n+1} \text{Bin}(2, d^{-i}) + U_{n+1}, \end{aligned}$$

where the last domination is due to the definition of U_{n+1} (see (33)). Advancing the induction completes the proof of (37).

The stochastic inequality (37) implies the existence of a coupling of U_n and U_n^s under which the inequality holds almost surely. All that remains is to show that this coupling is $(1, q)$ -bounded for the lower tail. For this, we observe that

$$\mathbb{P}\left(\sum_{i=1}^n \text{Bin}(2, d^{-i}) = 0\right) \geq \mathbb{P}\left(\sum_{i=1}^{\infty} \text{Bin}(2, d^{-i}) = 0\right) = \prod_{i=1}^{\infty} (1 - d^{-i})^2 \geq q.$$

Hence, under this coupling,

$$\mathbb{P}(U_n^s \leq U_n + 1 \mid U_n) \geq q.$$

□

Here in Lemma 5.17 and Proposition 5.18, we finally see how size-biasing creates a spine in $\vec{T}_{2d}$ along which there are extra visits to the root, as we described in Section 1.3. The $1 + \sum_{i=1}^n \text{Bin}(2, d^{-i})$ term in (37) is an overestimate of these extra returns. Heuristically, each $\text{Bin}(2, d^{-i})$ term represents two extra particles at level i on the spine. Because a particle has a d^{-1} chance of moving toward the root at each step, the chance of a particle at level i reaching the root is only d^{-i} , and we get only $O(1)$ expected extra visits to the root.

REMARK 5.19. *The random variables (U_n) were introduced because they could be proven to be log-concave. But computer investigations suggest that the random variables (W_n) themselves are log-concave, at least in the $p = 1/2$ case. If we could prove this, it would allow us to improve Theorem 1.5 to a result on density. The proof would go as follows. First, W_n would admit a size-bias coupling as in the previous proposition. Together with Proposition 5.1(a), this would show that $\mathbb{P}(W_n = 0) = O(1/n)$, from which it would follow that $\mathbb{P}(V_t = 0) = O(1/t)$. This implies that a B -particle at the root would survive for time t with probability $O(1/t)$, proving that the density of B -particles decays at rate $O(1/t)$. Since the density of A - and B -particles is the same, this would show that $\rho_t = O(1/t)$.*

The last ingredient is a technical lemma that bounds the growth of a sequence satisfying a certain recursive bound.

LEMMA 5.20. *Suppose a sequence μ_n satisfies*

$$(38) \quad \mu_{n+1} \leq \frac{\mu_n}{1 - e^{-q\mu_n}}$$

for some $q > 0$. Then there exists $C > 0$ such that $\mu_n \leq C \log n$ for all $n \geq 2$.

PROOF. We proceed by induction. Choose the constant $C \geq 2/q$ large enough that $\mu_2 \leq C \log 2$, establishing the base case of the induction. Now, we assume $\mu_n \leq C \log n$ and advance the induction. Since $x \mapsto x/(1 - e^{-qx})$ is increasing, the inductive hypothesis together with (38) and our assumption $C \geq 2/q$ yields

$$\mu_{n+1} \leq \frac{C \log n}{1 - n^{-2}}.$$

Hence in order to show $\mu_{n+1} \leq C \log(n+1)$, it suffices to show that

$$(39) \quad \frac{\log n}{1 - n^{-2}} \leq \log(n+1).$$

To prove this, start with the inequality $\log n \leq n - 1$ for $n \geq 1$. Dividing both sides by $n^2 - 1$ yields

$$\frac{\log n}{n^2 - 1} \leq \frac{1}{n + 1}.$$

Rewriting the left-hand side,

$$\frac{\log n}{1 - n^{-2}} - \log n \leq \frac{1}{n + 1}.$$

Then adding $\log n$ to both sides and applying the inequality $(\log n) + 1/(n+1) \leq \log(n+1)$ shows (39). This completes the induction and proves that $\mu_n \leq C \log n$ for all $n \geq 2$. $\square$

PROOF OF UPPER BOUNDS IN PROPOSITION 5.1. We will bound $\mathbb{E}U_n$, which bounds $\mathbb{E}W_n$ by Lemma 5.14. Let $\mu_n = \mathbb{E}U_n$. We start by applying Proposition 5.18 and Proposition 5.6(b) to deduce that

$$\mathbb{P}(U_n = 0) \leq e^{-q\mu_n/2},$$

where q is the constant from Proposition 5.18. Thus,

$$(40) \quad \mathbb{E}U_n^> = \frac{\mu_n}{1 - \mathbb{P}(U_n = 0)} \leq \frac{\mu_n}{1 - e^{-q\mu_n/2}}.$$

Now consider the critical case $p = 1/2$. If $X \geq 1$ a.s., then $\mathbb{E}[AX] = \mathbb{E}X$ by Lemma 5.5. Thus $\mu_{n+1} = \mathbb{E}[\mathcal{A}U_n^>] = \mathbb{E}U_n^>$, which with (40) shows that

$$\mu_{n+1} \leq \frac{\mu_n}{1 - e^{-q\mu_n/2}}.$$

It follows from Lemma 5.20 that $\mu_n \leq C \log n$ for some $C > 0$ and all $n \geq 2$. By Lemma 5.14, we have $\mathbb{E}W_n \leq \mu_n$. This completes the proof of the upper bound in part (a), the critical case.

We now prove the upper bound on $\lim_{n \rightarrow \infty} \mathbb{E}W_n$ in the $p < 1/2$ case. Since $\mathbb{E}W_n$ is an increasing sequence, we need only show that $\mathbb{E}W_n$ is bounded by $C \log(1/\epsilon)$ where $\epsilon = (1/2) - p$. By Lemma 5.5 and (40),

$$\mu_{n+1} = \mathbb{E}[\mathcal{A}U_n^>] = \mathbb{E}U_n^> - 2\epsilon \leq \frac{\mu_n}{1 - e^{-q\mu_n/2}} - 2\epsilon.$$

Let $\varphi(x) = x/(1 - e^{-qx/2}) - 2\epsilon$ for $x > 0$, and set $\varphi(0) = \lim_{x \rightarrow 0} \varphi(x) = \frac{1}{2q} - 2\epsilon$. Some calculus shows that the function $x \mapsto \varphi(x) - x$ is positive at $x = 0$, is strictly decreasing on $[0, \infty)$, and approaches a limit of -2ϵ as $x \rightarrow \infty$. Hence, φ has a unique fixed point x_0 . Since φ is strictly increasing, if $x < x_0$, then $\varphi(x) < \varphi(x_0) = x_0$. It follows from $q, \epsilon < 1/2$ that

$\varphi(1) - 1 > 0$; hence $x_0 > 1$. Thus $1 = \mu_0 < x_0$, and hence $\mu_n < x_0$ for all n . By Lemma 5.14, we have $\mathbb{E}W_n \leq x_0$ for all n . All that remains is to estimate x_0 . For $x = \frac{2}{q} \log(1/\epsilon)$,

$$\varphi(x) - x = \frac{2\epsilon^2}{q(1-\epsilon)} \log(1/\epsilon) - 2\epsilon,$$

which is negative when ϵ is small. Hence $x_0 < \frac{2}{q} \log(1/\epsilon)$ for all sufficiently small ϵ . $\square$

5.5. Lower bounds on root visits. The lower bounds in Proposition 5.1 are much simpler than the upper bounds. By Lemma 5.5, a lower bound on $\mathbb{E}W_n$ follows from anticoncentration estimates (i.e., lower bounds on $\mathbb{P}(W_n = 0)$). The idea of the proof is that even if we assume that W_n is as concentrated as possible—that its distribution is a point mass—then the distribution of $W_{n+1} \stackrel{d}{=} \mathcal{A}W_n$ is still nonconcentrated enough to yield the correct lower bound.

LEMMA 5.21. *For all $n \geq 0$ and $p \leq 1/2$,*

$$\mathbb{P}(W_{n+1} = 0) \geq 4^{-\mathbb{E}W_n - 1}.$$

PROOF. Define $X = \max(W_n, 1)$. By Lemma 5.4 we have $\mathcal{A}X \succeq \mathcal{A}W_n$. By Lemma 5.3,

$$\mathbb{P}(W_{n+1} = 0) \geq \mathbb{P}(\mathcal{A}X = 0).$$

Let $X_1, \dots, X_d$ be i.i.d. copies of X , and let $Y_1, \dots, Y_d$ be i.i.d. with $\mathbb{P}(Y_i = 1) = p$ and $\mathbb{P}(Y_i = -1) = 1 - p$. By the definition of $\mathcal{A}$ given in (29) and the fact that $X_i \geq 1$,

$$\begin{aligned} \mathbb{P}(\mathcal{A}X = 0) &= \mathbb{E} \left[(1 - 1/d)^{\sum_{i=1}^d (X_i + Y_i)} \right] = \prod_{i=1}^d \mathbb{E} \left[(1 - 1/d)^{X_i} \right] \mathbb{E} \left[(1 - 1/d)^{Y_i} \right] \\ &\geq (1 - 1/d)^{d\mathbb{E}X} \geq 4^{-\mathbb{E}W_n - 1}. \end{aligned}$$

From the first to the second line, we have applied Jensen's inequality to both expectations and then observed that $(1 - 1/d)^{\mathbb{E}Y_i} \geq 1$, since $\mathbb{E}Y_i \leq 0$ from our assumption $p \leq 1/2$. The last inequality holds because $\mathbb{E}X \leq \mathbb{E}W_n + 1$, and $(1 - 1/d)^d \geq 1/4$ for $d \geq 2$. $\square$

PROOF OF LOWER BOUNDS IN PROPOSITION 5.1. Let $\mu_n = \mathbb{E}W_n$. It follows from Lemmas 5.21 and 5.5 that for all $n \geq 0$ and $\epsilon = 1/2 - p \geq 0$,

$$\mu_{n+2} \geq \mu_n - 2\epsilon + (\epsilon + 1/2)4^{-\mu_n - 1}.$$

Since $x \mapsto x - 2\epsilon + (\epsilon + 1/2)4^{-x-1}$ is increasing for all $x \geq 0$,

$$(41) \quad \mu_n \geq x \implies \mu_{n+2} \geq x - 2\epsilon + (\epsilon + 1/2)4^{-x-1}.$$

Now, consider the $p = 1/2$ case. Choose $c \leq 1/16$, observing that it then holds for all $n \geq 1$ that

$$(42) \quad 2c \leq n^{1-c \log 4} / 8.$$

Also take c small enough that $\mu_2 \geq c \log 2$. Using this statement and $\mu_1 \geq 0$ as base cases, we prove

$$(43) \quad \mu_n \geq c \log n$$

by carrying out one induction for odd n and one for even n . Assuming (43), we apply (41) to get

$$\mu_{n+2} \geq c \log n + n^{-c \log 4} / 8 \geq c(\log n + 2/n),$$

applying (42) for the last inequality. Since $\log x$ is concave, we have $\log n + 2/n \geq \log(n + 2)$, advancing the induction. This completes the proof for the $p = 1/2$ case.

For $p < 1/2$, we already know from the upper bounds of Proposition 5.1 that $\mu := \lim_n \mu_n$ exists, since μ_n is increasing. Let

$$x_0 = -1 + \frac{1}{\log 4} \log \left(\frac{\epsilon + 1/2}{2\epsilon} \right),$$

which is the solution for x to the equation $x = x - 2\epsilon + (\epsilon + 1/2)4^{-x-1}$. We claim that $\mu \geq x_0$. Indeed, if not, then by taking μ_n sufficiently close to μ , we could conclude from (41) that $\mu_{n+2} > \mu$, which is a contradiction. $\square$

APPENDIX A: RANDOM WALK ESTIMATES

We start with some standard facts about the maximum displacement of random walk. The first is a moderate deviations estimate for a (possibly) biased continuous-time random walk.

LEMMA A.1. *Let S_t be a rate 1 continuous-time nearest-neighbor random walk on $\mathbb{Z}$ started at the origin with expected increment $\epsilon \in [0, 1]$. Let $M_t = \sup_{0 \leq s \leq t} S_t$. For all $0 \leq x \leq 2t$,*

$$\mathbb{P}(M_t \geq x + 2\epsilon t) \leq e^{-x^2/4t}.$$

PROOF. Applying Doob's martingale inequality to $S_t - \epsilon t$,

$$(44) \quad \mathbb{P}(M_t \geq x + 2\epsilon t) \leq \mathbb{P} \left(\sup_{0 \leq s \leq t} (S_s - \epsilon s) \geq x + \epsilon t \right) \leq \frac{\mathbb{E} \exp(\lambda(S_t - \epsilon t))}{\exp(\lambda(x + \epsilon t))}$$

for all λ . The moment generating function of a single random step is $\cosh \lambda + \epsilon \sinh \lambda$. Let $N \stackrel{d}{=} \text{Poi}(t)$ be the number of steps taken by the random walk up to time t . Recalling that $\mathbb{E}[z^N] = \exp(t(z - 1))$,

$$(45) \quad \mathbb{E} \exp(\lambda S_t) = \mathbb{E}(\cosh \lambda + \epsilon \sinh \lambda)^N = \exp(t(\cosh \lambda + \epsilon \sinh \lambda - 1)).$$

Applying (45) to (44) with $\lambda = x/t$ gives

$$(46) \quad \mathbb{P}(M_t \geq x + 2\epsilon t) \leq \exp \left(t \left(\cosh \frac{x}{t} + \epsilon \sinh \frac{x}{t} - 1 \right) - 2\epsilon x - \frac{x^2}{t} \right).$$

For $0 \leq \theta \leq 2$, one can confirm using Taylor series that $\cosh \theta - 1 \leq \frac{3}{4}\theta^2$ and $\sinh \theta \leq 2\theta$. Applying this to (46) together with our assumption $0 \leq x \leq 2t$ gives the claimed result. $\square$

LEMMA A.2. *Let S_t and M_t be as in Lemma A.1. For any $k \geq 1$ it holds that*

$$\mathbb{P}(M_t \geq t + k) \leq \exp \left(\frac{-k^2}{2(t + k)} \right).$$

PROOF. The displacement of the random walk is bounded by the total number of steps it takes in time t , a Poisson random variable with mean t . The estimate is then a standard tail bound for the Poisson (see [7, Chapter 2]). $\square$

We also need a quick estimate on the local time at the origin of a random walk after time t .

LEMMA A.3. *Let L_t be the time spent at the origin after t steps of a simple random walk $(S_s)_{s \leq t}$ started at the origin. It holds that $\mathbb{E}L_t \leq \sqrt{t}$.*

PROOF. Applying Tanaka's formula for the local time [26] gives $\mathbb{E}L_t = \mathbb{E}|S_t|$. It is a standard and straightforward recursion to show that the expected value of the square of a simple random walk after n steps is n . Hence, conditional on S_t taking $N \stackrel{d}{=} \text{Poi}(t)$ steps, we have

$$\mathbb{E}[|S_t| \mid N] \leq \sqrt{\mathbb{E}[S_t^2 \mid N]} = \sqrt{N}.$$

It follows by taking expectation and applying Jensen's inequality that

$$\mathbb{E}|S_t| \leq \mathbb{E}\sqrt{N} \leq \sqrt{\mathbb{E}N} = \sqrt{t}.$$

Hence $\mathbb{E}L_t \leq \sqrt{t}$. □

Our final goal is to prove the following result, which combines the gambler's ruin computation of the probability of a random walk hitting b before $-a$ with the moderate deviations bound given in Lemma A.1:

PROPOSITION A.4. *Let T_x be the hitting time of x for a continuous-time simple random walk on $\mathbb{Z}$ started at the origin. For any $t > 0$ and integers $a > 0$ and $3\sqrt{t} \leq x \leq 2t$,*

$$\mathbb{P}(T_x < T_{-a} \text{ and } T_x \leq t) \leq \frac{2a}{a+x} e^{-x^2/12t}.$$

Essentially, this result is that the moderate deviations result Lemma A.1 still holds (with a worse constant in the exponent) after conditioning the random walk to hit x before $-a$. Heuristically, we can see that this should be true by considering the analogous situation for Brownian motion. Conditioning the random walk to hit x before $-a$ is like conditioning Brownian motion to stay positive, which makes it a Bessel-3 process, the absolute value of a 3-dimensional Brownian motion. From this explicit representation, it is easy to check that it satisfies a moderate deviations tail bound.

To prove this result for random walks, we start by considering two random walks, one unbiased and one with a bias in the positive direction. We will need a lemma establishing the highly intuitive fact that the biased walk is faster to hit b even after conditioning both walks to hit b before $-a$, for any $a, b > 0$.

LEMMA A.5. *Let T_m be the first hitting time of m for a continuous-time simple random walk started from the origin, and let T'_m be the hitting time for a random walk that jumps to the right with probability $p > 1/2$. For any positive integers a and b , the conditional distribution of T_b given $T_b < T_{-a}$ is stochastically larger than that of T'_b given $T'_b < T'_{-a}$.*

PROOF. Let (S_n) be simple random walk conditioned to hit b before $-a$, and let (S'_n) be the biased walk under the same conditioning, both in discrete time. Take U and U' to be the hitting times of b for these processes. It suffices to prove $U' \preceq U$, since we can couple the continuous-time walks to follow the paths of the discrete-time walks with identical jump times. Let $\mathcal{S}_m$ denote the set of nearest-neighbor walks of length m from 0 to b that never hit $-a$, and note that both $(S_n)_{n=0}^U$ and $(S'_n)_{n=0}^{U'}$ are supported on the set $\cup_m \mathcal{S}_m$. Observe that each walk in $\mathcal{S}_m$ takes $r = (m+b)/2$ jumps to the right and $\ell = (m-b)/2$ jumps to the left. For any $(s_0, \dots, s_m) \in \mathcal{S}_m$,

$$\frac{\mathbb{P}\left((S'_n)_{n=0}^{U'} = (s_n)_{n=0}^m\right)}{\mathbb{P}\left((S_n)_{n=0}^U = (s_n)_{n=0}^m\right)} = \frac{p^r (1-p)^\ell / \mathbb{P}(T'_b < T'_{-a})}{2^{-m} / \mathbb{P}(T_b < T_{-a})} = h_0 (2\sqrt{p(1-p)})^m,$$

where

$$h_0 = \frac{p^{b/2} \mathbb{P}(T_b < T_{-a})}{(1-p)^{b/2} \mathbb{P}(T'_b < T'_{-a})}.$$

Thus,

$$\frac{\mathbb{P}(U' = m)}{\mathbb{P}(U = m)} = \frac{\sum_{(s_n) \in \mathcal{S}_m} \mathbb{P}\left((S'_n)_{n=0}^{U'} = (s_n)_{n=0}^m\right)}{\sum_{(s_n) \in \mathcal{S}_m} \mathbb{P}\left((S_n)_{n=0}^U = (s_n)_{n=0}^m\right)} = h_0 (2\sqrt{p(1-p)})^m.$$

This is decreasing in m , which proves $U' \preceq U$. $\square$

PROOF OF PROPOSITION A.4. The gambler's ruin calculation states that $\mathbb{P}(T_x < T_{-a}) = a/(a+x)$. Thus our goal is to prove that

$$(47) \quad \mathbb{P}(T_x \leq t \mid T_x < T_{-a}) \leq 2e^{-x^2/12t}.$$

To show this, fix t , and let (S_s) be a continuous-time random walk starting from 0 with probability $(1+t^{-1/2})/2$ of jumping to the right. By Lemma A.5, it suffices to bound its probability of hitting x in time t given that it hits x before $-a$. Since (S_s) must pass through $\lceil x/2 \rceil$ on its way to x and is a Markov process even after conditioning, it suffices to prove the bound under the assumption that (S_s) starts at $\lceil x/2 \rceil$ rather than the origin. Thus, our goal now is to show that

$$\mathbb{P}(U_x \leq t \mid U_x < U_{-a}) \leq 2e^{-x^2/12t},$$

where U_y is the first hitting time of y for the biased random walk starting at $\lceil x/2 \rceil$. By the gambler's ruin calculation for biased random walks,

$$\mathbb{P}(U_x < U_{-a}) = \frac{1 - \alpha^{\lceil x/2 \rceil + a}}{1 - \alpha^{x+a}},$$

where $\alpha = (1-t^{-1/2})/(1+t^{-1/2})$. Using $\alpha \leq 1-t^{-1/2}$ and our assumption $x \geq 3\sqrt{t}$,

$$\mathbb{P}(U_x < U_{-a}) \geq 1 - \alpha^{x/2+a} \geq 1 - (1-t^{-1/2})^{3\sqrt{t}/2+a} \geq 1 - e^{-3/2}.$$

Hence, by Lemma A.1,

$$\mathbb{P}(U_x \leq t \mid U_x < U_{-a}) \leq \frac{\mathbb{P}(U_x \leq t)}{\mathbb{P}(U_x < U_{-a})} \leq \frac{\exp\left(-\frac{x-2\sqrt{t}}{4t}\right)}{1 - e^{-3/2}} \leq 2e^{-x/12t},$$

using the bound $x - 2\sqrt{t} \geq x/3$ that follows from our assumption $x \geq 3\sqrt{t}$. This completes the proof by establishing (47). $\square$

Acknowledgments. We thank Michael Damron for helpful feedback and his assistance with Section 4.2.

Funding. Johnson was partially supported by the NSF-DMS Grant #1811952, Junge by the NSF-DMS Grant #185551, and Lyu by DMS-2206296 and DMS-2010035. Lyu and Sivakoff were partially supported by the NSF grant CCF-1740761.

REFERENCES

- [1] AHLBERG, D., GRIFFITHS, S. and JANSON, S. (2021). To fixate or not to fixate in two-type annihilating branching random walks. *Ann. Probab.* **49** 2637–2667. <https://doi.org/10.1214/21-aop1521> MR4317715
- [2] ARRATIA, R. (1981). Limiting point processes for rescalings of coalescing and annihilating random walks on $\mathbb{Z}^d$. *The Annals of Probability* 909–936.
- [3] ARRATIA, R. and BAXENDALE, P. (2015). Bounded size bias coupling: a Gamma function bound, and universal Dickman-function behavior. *Probab. Theory Related Fields* **162** 411–429. <https://doi.org/10.1007/s00440-009-0253-3>
- [4] ARRATIA, R., GOLDSTEIN, L. and KOCHMAN, F. (2019). Size bias for one and all. *Probab. Surv.* **16** 1–61. <https://doi.org/10.1214/13-ps221> MR3896143
- [5] AUFFINGER, A., DAMRON, M. and HANSON, J. (2017). *50 years of first-passage percolation* **68**. American Mathematical Soc.
- [6] BAHL, R., BARNET, P., JOHNSON, T. and JUNGE, M. (2021). Diffusion-limited annihilating systems and the increasing convex order. *arXiv preprint arxiv:2104.12797*.
- [7] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. OUP Oxford.
- [8] BRAMSON, M. and GRIFFEATH, D. (1980). Asymptotics for interacting particle systems on $\mathbb{Z}^d$. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **53** 183–196. <https://doi.org/10.1007/BF01013315>
- [9] BRAMSON, M. and LEBOWITZ, J. L. (1988). Asymptotic Behavior of Densities in Diffusion-Dominated Annihilation Reactions. *Phys. Rev. Lett.* **61** 2397–2400. <https://doi.org/10.1103/PhysRevLett.61.2397>
- [10] BRAMSON, M. and LEBOWITZ, J. L. (1991). Asymptotic behavior of densities for two-particle annihilating random walks. *J. Statist. Phys.* **62** 297–372. <https://doi.org/10.1007/BF01020872> MR1105266
- [11] BRENTI, F. (1989). Unimodal, log-concave and Pólya frequency sequences in combinatorics. *Mem. Amer. Math. Soc.* **81** viii+106. MR963833
- [12] CABEZAS, M., ROLLA, L. T. and SIDORAVICIUS, V. (2014). Non-equilibrium phase transitions: activated random walks at criticality. *J. Stat. Phys.* **155** 1112–1125. <https://doi.org/10.1007/s10955-013-0909-3> MR3207731
- [13] CABEZAS, M., ROLLA, L. T. and SIDORAVICIUS, V. (2018). Recurrence and density decay for diffusion-limited annihilating systems. *Probab. Theory Related Fields* **170** 587–615. <https://doi.org/10.1007/s00440-017-0763-3> MR3773795
- [14] CHEN, X., DERRIDA, B., HU, Y., LIFSHITS, M. and SHI, Z. (2019). A Max-Type Recursive Model: Some Properties and Open Questions. In *Sojourns in Probability Theory and Statistical Physics - III* (V. SIDORAVICIUS, ed.) 166–186. Springer Singapore, Singapore.
- [15] CHEN, X., DAGARD, V., DERRIDA, B., HU, Y., LIFSHITS, M. and SHI, Z. (2021). The Derrida-Retaux conjecture on recursive models. *Ann. Probab.* **49** 637–670. <https://doi.org/10.1214/20-aop1457> MR4255128
- [16] COOK, N., GOLDSTEIN, L. and JOHNSON, T. (2018). Size biased couplings and the spectral gap for random regular graphs. *Ann. Probab.* **46** 72–125. <https://doi.org/10.1214/17-AOP1180> MR3758727
- [17] CRISTALI, I., JIANG, Y., JUNGE, M., KASSEM, R., SIVAKOFF, D. and YORK, G. (2021). Two-type annihilating systems on the complete and star graph. *Stochastic Process. Appl.* **139** 321–342. <https://doi.org/10.1016/j.spa.2021.05.004> MR4275051
- [18] DAMRON, M., LYU, H. and SIVAKOFF, D. (2021). Stretched exponential decay for subcritical parking times on $\mathbb{Z}^d$. *Random Structures Algorithms* **59** 143–154. <https://doi.org/10.1002/rsa.21001> MR4284168
- [19] DAMRON, M., GRAVNER, J., JUNGE, M., LYU, H. and SIVAKOFF, D. (2019). Parking on transitive unimodular graphs. *The Annals of Applied Probability* **29** 2089–2113.
- [20] DAUVERGNE, D. and SLY, A. (2021). Spread of infections in a heterogeneous moving population. *arXiv preprint arXiv:2105.11947*.
- [21] DERRIDA, B. and RETAUX, M. (2014). The depinning transition in presence of disorder: a toy model. *J. Stat. Phys.* **156** 268–290. <https://doi.org/10.1007/s10955-014-1006-y>
- [22] GHOSH, S. and GOLDSTEIN, L. (2011). Concentration of measures via size-biased couplings. *Probab. Theory Related Fields* **149** 271–278. <https://doi.org/10.1007/s00440-009-0253-3> MR2773032 (2012e:60056)
- [23] GOLDSCHMIDT, C. and PRZYKUCKI, M. (2019). Parking on a random tree. *Combin. Probab. Comput.* **28** 23–45. <https://doi.org/10.1017/S0963548318000457> MR3917904
- [24] HARDY, R. and HARRIS, S. C. (2009). A spine approach to branching diffusions with applications to $\mathcal{L}^p$ -convergence of martingales. In *Séminaire de Probabilités XLII. Lecture Notes in Math.* **1979** 281–330. Springer, Berlin. MR2599214

- [25] HU, Y. and SHI, Z. (2018). The free energy in the Derrida-Retaux recursive model. *J. Stat. Phys.* **172** 718–741. <https://doi.org/10.1007/s10955-018-2066-1> MR3827300
- [26] KALLENBERG, O. (2006). *Foundations of modern probability*. Springer Science & Business Media.
- [27] KANG, K. and REDNER, S. (1984). Scaling Approach for the Kinetics of Recombination Processes. *Phys. Rev. Lett.* **52** 955–958. <https://doi.org/10.1103/PhysRevLett.52.955>
- [28] KOZA, Z. (1996). The long-time behavior of initially separated $A + B \rightarrow 0$ reaction-diffusion systems with arbitrary diffusion constants. *J. Statist. Phys.* **85** 179–191. <https://doi.org/10.1007/BF02175561> MR1413241
- [29] LACKNER, M.-L. and PANHOLZER, A. (2016). Parking functions for mappings. *J. Combin. Theory Ser. A* **142** 1–28. <https://doi.org/10.1016/j.jcta.2016.03.001> MR3499489
- [30] LEE, B. P. and CARDY, J. (1995). Renormalization group study of the $A + B \rightarrow \emptyset$ diffusion-limited reaction. *J. Statist. Phys.* **80** 971–1007. <https://doi.org/10.1007/BF02179861> MR1349773
- [31] LIGGETT, T. M. (1997). Ultra logconcave sequences and negative dependence. *J. Combin. Theory Ser. A* **79** 315–325. MR1462561
- [32] LYONS, R., PEMANTLE, R. and PERES, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164
- [33] LYONS, R. and PERES, Y. (2016). *Probability on trees and networks* **42**. Cambridge University Press.
- [34] OVCHINNIKOV, A. A. and ZELDOVICH, Y. B. (1978). Role of density fluctuations in bimolecular reaction kinetics. *Chem. Phys.* **28** 215–218.
- [35] PEKÖZ, E. A. and RÖLLIN, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** 587–608. <https://doi.org/10.1214/10-AOP559> MR2789507
- [36] PRZYKUCKI, M., ROBERTS, A. and SCOTT, A. (2019). Parking on the integers. *arXiv preprint arXiv:1907.09437*.
- [37] RIVERA, N., SAUERWALD, T., STAUFFER, A. and SYLVESTER, J. (2019). The dispersion time of random walks on finite graphs. In *The 31st ACM Symposium on Parallelism in Algorithms and Architectures* 103–113.
- [38] ROSS, N. (2011). Fundamentals of Stein’s method. *Probab. Surv.* **8** 210–293. <https://doi.org/10.1214/11-PS182> MR2861132
- [39] SHAKED, M. and SHANTHIKUMAR, J. G. (2007). *Stochastic orders*. Springer Series in Statistics. Springer, New York. <https://doi.org/10.1007/978-0-387-34675-5> MR2265633 (2008g:60005)
- [40] SHI, Z. (2015). *Branching random walks*. *Lecture Notes in Mathematics* **2151**. Springer, Cham Lecture notes from the 42nd Probability Summer School held in Saint Flour, 2012, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. <https://doi.org/10.1007/978-3-319-25372-5> MR3444654
- [41] STANLEY, R. P. (1989). Log-concave and unimodal sequences in algebra, combinatorics, and geometry. In *Graph theory and its applications: East and West (Jinan, 1986)*. *Ann. New York Acad. Sci.* **576** 500–535. New York Acad. Sci., New York. MR1110850
- [42] TOUSSAINT, D. and WILCZEK, F. (1983). Particle–antiparticle annihilation in diffusive motion. *J. Chem. Phys.* **78** 2642–2647.