The Dispersion Process Has the Same Phase Transition on Almost Every Graph

Julius Hallmann
Department of Mathematics, LMU Munich, Germany

Kostas Lakis
Department of Computer Science, ETH Zürich, Switzerland

Tamás Makai
Department of Mathematics, LMU Munich, Germany

Abstract

The *Dispersion process* was introduced by Cooper, McDowell, Radzik, Rivera and Shiraga (2018), in which a number of particles are initially placed on a given vertex of a graph and update their positions according to the following dynamics. In each round, the particles which are not alone on a vertex (called *unhappy*) simultaneously move to a uniformly random neighbor. In contrast, the rest of the particles (called *happy*) stay put. The process terminates once every particle is happy.

When the process runs on the complete graph, they showed that there is a phase transition with respect to the running time when the number of particles reaches $n/2$. Below this threshold the running time is logarithmic, while above the threshold it is exponential. We show that the same behavior holds for the binomial random graph $G(n, 1/2)$ with high probability.

The main difference to the complete graph is that the number of happy particles in the neighborhood of a vertex can vary significantly from vertex to vertex, resulting in differing local behaviors. Fortunately the number of such vertices is limited and thus they only have a negligible effect on the running time of the process.

2012 ACM Subject Classification Mathematics of computing → Markov processes; Mathematics of computing → Random graphs

Keywords and phrases Dispersion process, Random Graphs, Drift analysis

Digital Object Identifier 10.4230/LIPIcs.AofA.2026.25

Funding *Kostas Lakis*: Gratefully acknowledges support by the Swiss National Science Foundation [grant number 10003390].

1 Introduction

The *Dispersion process* introduced by Cooper, McDowell, Radzik, Rivera and Shiraga in 2018 [1] is a synchronous process that proceeds in discrete time-steps. Initially $M \geq 2$ indistinguishable particles are placed on a given vertex v of a connected graph $G = (V, E)$. A particle is called *happy* if no other particle occupies the same vertex it sits on and *unhappy* otherwise. In each round the unhappy particles simultaneously jump to a neighbor chosen uniformly and independently at random while the happy particles stay where they are. The process stops at the first round in which each particle is happy. This (random) time is called *dispersion time* and denoted by $T_{v,G,M}$. When the vertex of origin is clear from the context, we may drop the subscript v.

Intuitively, dispersion should happen faster if there are fewer particles, as collisions become less likely. When G is the complete graph, there is a phase transition when the number of particles reaches $n/2$, where the dispersion time switches from logarithmic to exponential.

▶ **Theorem 1** (Theorem 1 in [1]). *Let* $0 < \delta < 1$.
1. *If* $M = \frac{1-\delta}{2}n$, *then* $T_{K_n,M} = \mathcal{O}(\log n)$ *with high probability*[1].
2. *If* $M = \frac{1+\delta}{2}n$, *then* $\log(T_{K_n,M}) = \Omega(n)$ *with high probability.*

The behavior of the process around the threshold was refined in subsequent works. In [2] it was shown that the phase transition is smooth, and a critical window of size $\delta = \mathcal{O}(n^{-1/2})$ was established within which $T_{K_n,M} = \Theta(n^{1/2})$ typically. Finally [3] analyzed the running time of the process within the critical window. In particular it was shown that the dispersion time normalized by $n^{-1/2}$ converges in distribution to a continuous random variable that only depends on $\lim_{n\to\infty} \delta n^{1/2}$.

Besides the complete graph the behavior of the dispersion process was studied on a wide range of graphs in the original paper [1]. These include, among others, the infinite line, the d-regular infinite trees, lattices and hypercubes. In addition to the dispersion time they also studied the *dispersion distance*, i.e. the maximal distance of a particle from the origin at the end of the process.

For the infinite path it was shown that typically the dispersion distance is between $M/2$ and $\mathcal{O}(M \log M)$ with high probability, which was later sharpened by Frieze and Pegden [4] to $\Theta(M)$. Later Shang [7] extended their results on the dispersion distance to a variation of the dispersion process on the line, where the jump destinations are nonuniform.

This work is concerned with dispersion on the random graph $G(n, 1/2)$. More precisely, we first place M particles on vertex v_0 for some $v_0 \in V$ then sample a graph $G \sim G(n, 1/2)$ and afterwards perform the dispersion process on G. Our main result is that in this regime the dispersion time behaves essentially the same as on the complete graph.

Model

Let $n \in \mathbb{N}$ and denote by $M \geq 2$ the number of particles. For $i \in [M]$ and $t \in \mathbb{N}_0$, we write $p_{i,t}$ for the position of the i-th particle at time t. Initially, we have $p_{i,0} = v_0$ for all $i \in [M]$ for some $v_0 \in V(G)$. Then, in distribution,

$$p_{i,t+1} = \begin{cases} p_{i,t}, & p_{i,t} \neq p_{j,t} \ \ \forall j \in [M] \setminus \{i\} \\ X_{i,t,p_{i,t}}, & \text{otherwise} \end{cases},$$

where $\{X_{i,t,v} : i \in [M], t \in \mathbb{N}_0, v \in V\}$ are independent random variables such that $X_{i,t,v}$ has uniform distribution on $N(v)$ and $N(v)$ denotes the neighborhood of v in G. Denote by $\mathcal{P}(t) = \{p_{i,t} : i \in [M]\}$.

From this we derive some auxiliary variables. The amount of particles on a vertex $v \in V$ at time $t \in \mathbb{N}_0$ is $a_v(t) = |\{i \in [M] : p_{i,t} = v\}|$. We denote the set of vertices occupied by unhappy particles as $U(t) = \{v \in V : a_v(t) > 1\}$ and the number of unhappy particles as $U_p(t) = \sum_{v\in U(t)} a_v(t)$. The set of vertices occupied by happy particles is $H(t) = \{v \in V : a_v(t) = 1\}$ and since every vertex can contain at most 1 happy particle the number of happy particles at time t is $|H(t)|$. Our main quantity of interest, the dispersion time, is the first time each particle is happy, that is

$$T_{G(n,1/2),M} = \inf\{t \geq 0 : |U(t)| = 0\}.$$

▶ **Theorem 2.** *Let* $0 < \delta < 1$.
1. *If* $M = \frac{1-\delta}{2}n$, *then* $T_{G(n,1/2),M} = \mathcal{O}(\log n)$ *with high probability;*
2. *If* $M = \frac{1+\delta}{2}n$, *then* $\log(T_{G(n,1/2),M}) = \Omega(n)$ *with high probability.*

[1] An event occurs with high probability, if its probability tends to 1 as $n \to \infty$

Outline

We give here a brief intuitive outline of the arguments in our proofs. For the sake of exposition, let us first explain how one can show the phase transition in the easier case of the complete graph with self-loops. In such a scenario, each unhappy particle chooses a uniformly random vertex of the graph to jump to in the next round. If there are currently hn happy particles, then with probability h a fixed unhappy particle hits one of them. This collision will lead to 2 unhappy particles in the next round. Therefore, if $h > 1/2$, in expectation we have more than 1 unhappy particles being "born" out of a single unhappy particle.

Of course, there could be other reasons for a particle to end up unhappy (e.g. some other unhappy particle happens to choose the same vertex to jump to). However, when the number of unhappy particles is, say $o(\sqrt{n})$, this effect is negligible, as such collisions are extremely unlikely. Therefore, in this terminal part of the process, the number of unhappy particles resembles a branching process where each individual either dies or gives birth to two new individuals with probability $1 - h$ and h, respectively. Since towards the end we have h being essentially equal to M/n, the reason for the phase transition at $M = n/2$ is now evident, as this decides whether the expected progeny of a single individual in the branching process is higher or lower than 1.

In the subcritical regime we use drift analysis to make this connection formal. Instead of considering the number of unhappy particles (as has been done in previous works), we focus on the number of pairs of particles which are sitting on the same vertex, i.e. we look at the potential function

$$\sum_{v} \binom{a_v}{2},$$

with the usual convention that $\binom{1}{2} = 0$. This potential function reaches 0 exactly when no more unhappy particles are left. However it turns out easier to analyse than the number of unhappy particles, especially in the case of a random graph.

For the supercritical regime, we predominantly rely on the method of bounded differences to establish that the number of unhappy particles remains above a certain value for some exponential number of steps.

The extra difficulty arising from the random graph setting is that even though most vertices behave very regularly with respect to the numbers of edges towards most sets of other vertices, there are still some combinations where the number of edges is much higher or lower than the expectation. For example, by definition a vertex v is connected to every vertex in its neighborhood (comprised of the set $S = \{u_1, \ldots, u_{d(v)}\}$) while a priori the expected number of edges from v to S is only $|S|/2$. This poses a problem for the analysis of the drift in situations where for example the happy particles are somehow all concentrated in the neighborhood of a single vertex. This implies that any unhappy particle sitting on such a vertex will deterministically hit a happy particle in the next round, deviating from the picture we saw for the complete graph.

One could attempt to analyze the complicated random walks of the particles and argue that such a case simply does not occur. However, we take a different approach which we believe is simpler. In particular, we accept the fact that such situations may in principle occur, but control the number of vertices which could simultaneously exhibit such a strong connection to the set of vertices with happy particles on them. We may refer to such a vertex (i.e. one with only one particle on it) as a *happy vertex*. In the case under investigation, one can actually show that the following pseudorandom property holds. For any set W, the number of vertices which have too many or too few edges (a linear number more or fewer than the expectation) to W is at most a constant. Using this pseudorandom property, we can deal with the complications introduced above. For the subcritical regime, we slightly modify our potential function to punish configurations which put unhappy particles on vertices whose neighborhoods contain substantially more happy vertices than expected. We call such vertices *bad*. On the other hand in the supercritical regime we show that only a tiny portion of the unhappy particles are on vertices with substantially *fewer* happy vertices in its neighborhood, which have an insignificant effect on the outcome.

Preliminaries

Below we collect some facts which will be useful for our analysis. We start with a theorem that we use to establish the logarithmic running time.

▶ **Theorem 3** (Theorem 18 [6]). *Let $(X_t)_{t\geq 0}$ be a sequence of non-negative random variables with a finite state space $\mathcal{S} \subseteq \mathbb{R}_0^+$ such that $0 \in \mathcal{S}$. Let $s_{min} := \min(\mathcal{S} \setminus \{0\})$, and let $T := \inf\{t \geq 0 \mid X_t = 0\}$. Suppose that $X_0 = s_0$, and that there exists a $\gamma > 0$ such that for all $s \in \mathcal{S} \setminus \{0\}$ and all $t \geq 0$,*

$$\mathbb{E}[X_t - X_{t+1} \mid X_t = s] \geq \gamma s.$$

Then, for all $r \geq 0$,

$$\mathbb{P}\left[T > \left\lceil \frac{r + \ln(s_0/s_{min})}{\gamma} \right\rceil\right] \leq e^{-r}.$$

Next we recall several versions of the Chernoff bound for binomial random variables used throughout the paper; see, e.g., Section 2.1 in [5].

▶ **Theorem 4.** *If $X \sim Bin(n,p)$ and $\lambda = np$, then for $t \geq 0$*

$$\begin{aligned}
\mathbb{P}[X \geq \mathbb{E}[X] + t] &\leq \exp\left(-\lambda((1+t/\lambda)\log(1+t/\lambda) - t/\lambda)\right), \\
\mathbb{P}[X \leq \mathbb{E}[X] - t] &\leq \exp\left(-\lambda((1-t/\lambda)\log(1-t/\lambda) + t/\lambda)\right), \\
\mathbb{P}[X \geq \mathbb{E}[X] + t] &\leq \exp\left(-t^2/(2(\lambda + t/3))\right), \\
\mathbb{P}[X \leq \mathbb{E}[X] - t] &\leq \exp\left(-t^2/(2\lambda)\right).
\end{aligned}$$

In the supercritical regime we will make use of the method of bounded differences.

▶ **Theorem 5** (Corollary 2.27 [5]). *Let $X_1, \ldots, X_n$ be independent random variables where X_i is $\mathcal{X}_i$-valued for some set $\mathcal{X}_i$, and let $X = (X_1, \ldots, X_n)$. Assume $f : \mathcal{X}_i \times \cdots \times \mathcal{X}_n \to \mathbb{R}$ is a measurable function with the following Lipschitz property. There exist $c_i > 0$, $i \in [n]$ such that if $x, x' \in \prod_{i\in[n]} \mathcal{X}_i$ differ only in the i-th coordinate, then $|f(x) - f(x')| \leq c_i$. Then, for $t \geq 0$,*

$$\mathbb{P}[|f(X) - \mathbb{E}[f(X)]| \geq t] \leq 2\exp\left(-\frac{t^2}{2\sum_{i\in[n]} c_i^2}\right).$$

Layout of the paper

In Section 2 we provide the pseudorandom properties of $G(n, 1/2)$ that will be useful for our analysis. We treat the subcritical and supercritical regimes in Section 3 and Section 4, respectively.

2 Pseudorandom properties

In this section we collect a set of pseudorandom properties that hold with high probability in $G(n,1/2)$ and that we will use in our analysis of the process. We start with a standard degree and codegree concentration result for $G(n,1/2)$.

▶ **Lemma 6.** *With high probability,*

$$\max_{v\in V(G)} \left|d(v)-n/2\right| \le C\sqrt{n\log n}, \quad \max_{u,v\in V(G)} \left|\mathrm{codeg}(u,v)-n/4\right| \le C\sqrt{n\log n}$$

for an absolute constant $C>0$.

In our analysis of the drift, the situation looks fairly similar to the case of the complete graph for most vertices, but there are some blemishes that we have to control. Most of these are in the form of vertices that have a disproportionate number of neighbours in certain sets, for example the set of happy vertices. The following lemma quantifies the prevalence of such vertices.

For a set $W\subset V$ and for all $\beta>0$, let $B_\beta(W)$ denote the set of vertices $v\in V$ such that $|d_W(v)-|W|/2|\ge \beta n$, where $d_W(v)$ denotes the number of neighbors of v in W.

▶ **Lemma 7.** *For any $\beta>0$ with high probability for all $W\subset V$ we have $|B_\beta(W)|\le 6/\beta^2$.*

Proof. Without loss of generality assume that $\beta\le 1$. Fix W and a set U of size $k=\lfloor 6/\beta^2\rfloor+1$. For $u\in U$ let $\neg\mathcal{E}_u$ be the event that $|W|/2-\beta n\le d_{W\setminus U}(u)\le |W|/2+\beta n-k$ and note that for any vertex $u\in B_\beta(W)$ we have $\mathcal{E}_u$. The Chernoff bound implies that $\mathbb{P}[\mathcal{E}_u]\le 2\exp(-\beta^2 n/3)$.

Note that the events $\mathcal{E}_u$ are independent for $u\in U$ and thus the probability of there existing a set U with $|U|=k$ such that $\mathcal{E}_u$ holds for every $u\in U$ is at most

$$\binom{n}{k}2^k\exp(-k\beta^2n/3)\le \left(2n\exp(-\beta^2n/3)\right)^k.$$

This in turn implies $\mathbb{P}[|B_\beta(W)|\ge k]\le \left(2n\exp(-\beta^2 n/3)\right)^k$ and by applying the union bound the probability that there exists a set W with $|B_\beta(W)|\ge 6/\beta^2$ is at most

$$2^n\left(2n\exp(-\beta^2n/3)\right)^{6/\beta^2}=o(1),$$

completing the proof. ◀

3 The subcritical regime

In this section we will prove part 1) of Theorem 2. Fix δ and a sufficiently small β with respect to δ. In the following, we will assume that our graph has every property described in Lemma 6, and the statement of Lemma 7 holds for both β and $\beta/2$.

Our first goal is to show that after the first step the maximum number of particles on a single vertex is at most $\log n$. This polylogarithmic bound on the maximum "load" on each vertex will be useful for later arguments.

For the following, let L_t be this maximum number at time t. In particular, $L_0=M$. After one step these particles will disperse in such a way that any vertex will end up containing at most $\log n$ particles. Our aim is to show that with high probability $L_i\le \log n$ for $1\le i\le C\log n$ for any $C>0$.

▶ **Lemma 8.** *Let $C > 0$. With high probability for all $1 \leq i \leq C \log n$ we have*

$$L_i \leq \log n.$$

Proof. Recall from Lemma 6 that $|d(v) - n/2| \leq \eta n$ for some $\eta = \mathcal{O}(\sqrt{\log n/n})$. Note that for $i > 0$ the number of particles on any vertex after step i is stochastically dominated by $1 + \mathrm{Bin}(M, ((1/2 - \eta)n)^{-1})$ and thus by the Chernoff bound

$$\mathbb{P}[a_v(i) \geq \log n] \leq \exp(-1.1 \log n).$$

Applying the union bound over the n vertices and the $C \log n$ steps gives $L_i \leq \log n$ with high probability. ◀

We now come to the main part of our drift analysis.

$$\Phi(t) = 4 \sum_{v \in B_\beta(H(t))} \binom{a_v(t)}{2} + \sum_{v \in V \setminus B_\beta(H(t))} \binom{a_v(t)}{2}.$$

This potential counts every pair of unhappy particles on the same vertex, and in addition punishes unhappy particles more when they are placed on vertices with a disproportionate number of neighbors in $H(t)$, in particular we want to punish vertices which have too many happy neighbors, as the particles on these vertices are less likely to become happy in the following step.[2] However the potential will still decrease; a particle in $B_\beta(H(t))$ is much more likely to land on a vertex outside of $B_\beta(H(t))$, thereby reducing its contribution to the potential, while the particles outside of $B_\beta(H(t))$ have a good chance of becoming happy. The only extra issue with this potential is that the set of vertices $B_\beta(H(t))$ depends on t, which in principle could lead to a significant increase in the potential. However, since only a logarithmic number of particles can be found on a single vertex, this behavior only becomes relevant towards the end of the process, when there are only few unhappy particles left and these can cause only minor changes in the set of happy vertices.

Our goal is to show a multiplicative drift in the potential.

▶ **Lemma 9.** *Let $\mathcal{P}(t)$ satisfy $L_t \leq \log n$. Then $\mathbb{E}[\Phi(t+1) \mid \mathcal{P}(t)] \leq (1-\gamma)\Phi(t)$ for some $\gamma > 0$.*

Proof. By Lemma 6 we have for every $u, v \in V$ that $|d(v) - n/2| \leq \eta n$ and $|\mathrm{codeg}(u,v) - n/4| \leq \eta n$ for any $\eta > 0$. We will pick η to be sufficiently small with respect to δ and β (recall that the number of particles M is assumed to be $n(1-\delta)/2$). Note that

$$\mathbb{E}[\Phi(t+1) \mid \mathcal{P}(t)] \leq \sum_{u \in U(t)} \binom{a_u(t)}{2} \frac{1}{d(u)} + \frac{1}{2} \sum_{\substack{u,v \in U(t) \\ u \neq v}} \frac{\mathrm{codeg}(u,v) a_u(t) a_v(t)}{d(u)d(v)}$$
$$+ \sum_{u \in U(t)} \frac{a_u(t) d_{H(t)}(u)}{d(u)} + \mathcal{O}(1).$$

as there are $\mathcal{O}(1)$ vertices in $B_\beta(H(t+1))$ and the number of particles on any vertex is stochastically dominated by $1 + \mathrm{Bin}(M, ((1/2-\eta)n)^{-1})$ implying

$$\mathbb{E}\left[\binom{a_v(t+1)}{2} \mid \mathcal{P}(t)\right] = \mathcal{O}(1).$$

[2] The factor 4 is chosen so that some arguments are easier. In principle one could choose any number larger than 2 (intuitively because a particle on a bad vertex is deterministically assumed to hit a happy particle in the next round), but there is no real benefit and only technical overhead in doing so.

More precisely, all the contributions from particles jumping to bad vertices (with respect to the happy vertices at time $t+1$) are included in the $\mathcal{O}(1)$ term. We will first consider the case where $\Phi(t) \geq \log^2 n$. The first summand in our bound for $\mathbb{E}[\Phi(t+1) \mid \mathcal{P}(t)]$ is clearly $o(\Phi(t))$. Since $B_\beta(H(t)) = \mathcal{O}(1)$ and $L_t = \mathcal{O}(\log n)$ the third summand can be bounded from above by

$$\begin{aligned}\frac{U_p(t)(|H(t)|+2\beta n)}{2(1/2-\eta)n} + \mathcal{O}(\log n) &\leq 2(1+4\eta)(|H(t)|/n+2\beta)\frac{U_p(t)}{2} + \mathcal{O}(\log n)\\ &\leq 2(1+5\eta)(|H(t)|/n+2\beta)\Phi(t),\end{aligned}$$

as $U_p(t) \leq 2\Phi(t)$. Finally the middle term is bounded by

$$\frac{U_p(t)^2}{2}\frac{(1/4+\eta)}{(1/2-\eta)^2 n} \leq \frac{2U_p(t)}{n}\Phi(t).$$

Therefore

$$\mathbb{E}[\Phi(t+1) \mid \mathcal{P}(t)] \leq \left(\frac{2(1+5\eta)|H_t| + 2U_p(t)}{n} + 4(1+5\eta)\beta + o(1)\right)\Phi(t) \leq (1-\gamma_1)\Phi(t).$$

for some $\gamma_1 > 0$ where in the last step we used that $|H(t)| + U_p(t) = M \leq (1-\delta)n/2$.

Now consider the case in which $\Phi(t) < \log^2 n$ and note that in this case we have $U_p(t) < 2\log^2 n$. Note that $|H(t+1) \triangle H(t)| < 2\log^2 n$, implying that $B_\beta(H(t+1)) \subseteq B_{\beta/2}(H(t))$. Our aim is to establish an upper bound on the expected value of

$$4\sum_{v\in B_{\beta/2}(H(t))}\binom{a_v(t+1)}{2} + \sum_{v\in V\setminus B_{\beta/2}(H(t))}\binom{a_v(t+1)}{2}.$$

First we consider the contribution to the potential when two particles from the same vertex v jump to the same vertex. This is at most

$$\sum_{u\in U(t)}\binom{a_u(t)}{2}\frac{1}{d(u)} + 4\frac{|B_{\beta/2}(H(t))|}{((1/2-\eta)n)^2}\Phi(t) = o(\Phi(t)).$$

Next we consider 2 particles stemming from different vertices landing on the same vertex, which contributes at most

$$\frac{1}{2}\sum_{\substack{u,v\in U(t)\\ u\neq v}}\frac{\operatorname{codeg}(u,v)a_u(t)a_v(t)}{d(u)d(v)} + 4\binom{U_p(t)}{2}\frac{|B_{\beta/2}(H(t))|}{((1/2-\eta)n)^2} \leq \frac{U_p(t)^2}{n}.$$

Now we consider the particles starting from a vertex in $U(t) \setminus B_\beta(H(t))$ landing on a happy particle, which is at most

$$U_p(t)\frac{|H(t)|/2+\beta n}{(1/2-\eta)n} + 4U_p(t)\frac{|B_{\beta/2}(H(t))|}{(1/2-\eta)n} \leq (1+4\eta)(|H(t)|/n+2\beta)U_p(t).$$

Denote by $U_p'(t)$ the number of particles on vertices in $U(t)\cap B_\beta(H(t))$. Then the contribution from vertices starting in $B_\beta(H(t))$ and landing in $H(t)$ is at most $U_p'(t)(1+o(1))$. So for some sufficiently small $1/3 > \gamma_2 > 0$ we have an overall expectation

$$\begin{aligned}&o(\Phi(t)) + \frac{U_p(t)^2}{n} + (1+4\eta)\left(\frac{|H(t)|}{n} + 2\beta\right)U_p(t) + (1+o(1))U_p'(t)\\ &\quad\leq o(\Phi(t)) + \frac{1-\gamma_2}{2}U_p(t) + (1+o(1))U_p'(t).\end{aligned}$$

Now note that $(U_p(t) - U_p'(t)) + 4U_p'(t) = U_p(t) + 3U_p'(t) \leq 2\Phi(t)$ and thus

$$\frac{1-\gamma_2}{2}U_p(t) + (1+o(1))U_p'(t) \leq \frac{1-\gamma_2}{2}(U_p(t) + 3U_p'(t)) \leq (1-\gamma_2)\Phi(t),$$

and the result follows. ◀

We show that in the subcritical regime the dispersion time is logarithmic using Theorem 3.

Proof of Theorem 2 1). Let $\mathcal{E}(t)$ be the event that $L_i \leq (1+o(1))\log^4 n$ for $1 \leq i \leq t-1$. Consider $\Psi(t) = \Phi(t)\mathbb{1}[\mathcal{E}(t)]$. By Lemma 9 we have that $\mathbb{E}[\Psi(t+1) \mid \Psi(t)] \leq (1-\gamma)\Psi(t)$ and the result follows from Theorem 3 and Lemma 8. ◀

4 The supercritical regime

Fix $\delta > 0$ and a sufficiently small $\beta, \eta > 0$ with respect to δ. Similarly to the subcritical regime, we will need an upper bound on the number of particles on any given vertex. However, as the process runs for exponential time, we can no longer ensure that this bound is logarithmic.

▶ **Lemma 10.** *The probability that in step $t \geq 1$ there is a vertex with more than $n/\log n$ particles on it, is at most $e^{-n/2}$ for large enough n.*

Proof. By Lemma 6 we have that $(1/2-\eta)n \leq d(v) \leq (1/2+\eta)n$ for every $v \in V$ and $\eta \in \mathbb{R}$. Therefore the number of particles on v is stochastically dominated by $1+\text{Bin}(n, ((1/2-\eta)n)^{-1})$ and by Theorem 4 we have for sufficiently small η that $\mathbb{P}\left[\text{Bin}\left(n, ((1/2-\eta)n)^{-1}\right) \geq \frac{n}{\log n} - 1\right]$ is at most

$$\exp\left(-\frac{n}{\log n}\log\left(\frac{n}{3\log n}\right) + \frac{n}{\log n}\right),$$

where we upper bounded the mean of the binomial by 3. This is equal to $e^{-(1+o(1))n}$, and thus the result follows by taking the union bound over all vertices. ◀

The idea behind our argument is the following. Once there are only a few unhappy particles left we can essentially ignore collisions between unhappy particles. With this assumption, an unhappy particle jumping on a happy vertex results in two unhappy particles. The probability of this happening is close to $M/n > 1/2$, which gives us the desired drift, see Lemma 12. Now, this approach comes with two limitations. First, the number of unhappy particles must be small for our argument to work. Second, we want to show that $U_p(t)$ stays above a threshold for a very long (exponential) time. For this to hold we need strong concentration results, which in turn require the number of unhappy particles to be $\Omega(n)$. In addition, this has the benefit of bad vertices being irrelevant for the dynamics of the process. So, $U_p(t)$ is concentrated around a positive drift only within a window of size linear in n. The next lemma ensures that our process does not skip this window.

▶ **Lemma 11.** *Let $\mathcal{P}(t)$ satisfy $L_t \leq n/\log n$. For $U_p(t) = \Omega(n)$ we have*

$$\mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] \geq U_p(t)/1000.$$

Proof. We bound the expectation in two ways, suited for many and few unhappy particles. First, count the number of unhappy particles jumping to vertices occupied by happy particles. Assume $|H(t)| \geq n/4$. By assumption each vertex carries at most $n/\log n$ particles and the number of vertices with an unusually small share of happy particles in their neighborhood is constant by Lemma 7 (note that $|H(t)| = \Theta(n)$ in this case). Hence there are at least $(1 - o(1))U_p(t)$ unhappy particles on vertices v such that $|N(v) \cap H(t)| \geq (1-\beta)|H(t)|/2$ for any $\beta > 0$. Each of these particles jumps to a happy vertex with probability at least $(1-\beta)|H(t)|((1/2+\eta)n)^{-1}$ and remains unhappy. Then, if n is large enough

$$\mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] \geq (1-o(1))U_p(t)\frac{(1-\beta)|H(t)|}{(1+2\eta)n} \geq \frac{U_p(t)}{1000}.$$

Alternatively, we have $U_p(t) > n/4$. In this case we provide an upper bound on the probability that a particle on vertex v becomes happy. In order for this to happen it has to jump to an empty neighbor and no other particle may jump on it. This probability can be bounded from above by

$$\frac{1}{d(v)} \sum_{u \in N(v)} \left(1 - \frac{1}{(1/2+\eta)n}\right)^{-1+\sum_{w \in N(u) \cap U(t)} a_w(t)}.$$

Note that

$$\sum_{u \in N(v)} \sum_{w \in N(u) \cap U(t)} a_w(t) \geq (1/4 - \eta)n(1-o(1))U_p(t) \geq \frac{U_p(t)n}{5},$$

as every particle corresponding to $a_w(t)$ is counted at least $\mathrm{codeg}(v,w) \geq (1/4-\eta)n$ times.

Therefore, for at least one tenth of the $u \in N(v)$ we have that $\sum_{w \in N(u)} a_w(t) \geq U_p(t)/10$ as otherwise

$$\frac{1}{10}d(v)U_p(t) + \frac{9}{10}d(v)\frac{U_p(t)}{10} \leq \frac{19}{100}(1/2+\eta)nU_p(t) < \frac{U_p(t)n}{5},$$

which would be a contradiction to our earlier lower bound. This then leads to

$$\frac{1}{d(v)} \sum_{u \in N(v)} \left(1 - \frac{1}{(1/2+\eta)n}\right)^{-1+\sum_{w \in N(u)} a_w(t)} \leq \frac{9}{10} + \frac{1}{10}\exp\left(-\frac{U_p(t)}{(5+10\eta)n}\right) < 1 - 10^{-3},$$

and the proof is complete. ◀

In the following lemma we establish the positive drift in the number of unhappy particles.

▶ **Lemma 12.** *Let $\mathcal{P}(t)$ satisfy $L_t \leq n/\log n$. Assume $U_p(t)/n < \delta/8$ and $U_p(t) = \Omega(n)$. Then,*

$$\mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] \geq U_p(t)(1+\delta/3).$$

Proof. Since $|H(t)| = \Theta(n)$, the same reasoning as in Lemma 11 gives that the expected number of unhappy particles jumping to vertices occupied by happy particles is at least

$$(1-o(1))U_p(t)\frac{(1-\beta)(M - U_p(t))}{(1+2\eta)n}.$$

If an unhappy particle wakes up a happy particle there are two unhappy particles, so if there are only a few unhappy particles left (such that collisions become unlikely), the above lower bound becomes weak. Instead, we count the wake-ups as if no collisions occur and subtract the amount of those unhappy particles that jump to the same vertex as any other unhappy particle.

Fix two unhappy particles. The probability that they jump to the same vertex is at most $2((1-2\eta)n)^{-1}$. Hence, the probability of any unhappy particle jumping on the same vertex as the first unhappy particle is at most $2U_p(t)((1-2\eta)n)^{-1}$ using a union bound. Therefore,

$$\begin{aligned}\mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] &\geq (1-o(1))2U_p(t)\frac{(1-\beta)(M-U_p(t))}{(1+2\eta)n} - 2\frac{U_p(t)^2}{(1-2\eta)n} \\ &\geq U_p(t)\left(\frac{(1-\beta)(1+\delta-\delta/4)}{1+2\eta} - \frac{\delta/4}{1-2\eta}\right) \geq U_p(t)(1+\delta/3),\end{aligned}$$

which is the bound we wanted to show. ◀

We use the method of bounded differences to show that $U_p(t)$ is concentrated.

▶ **Lemma 13.** *For every $\gamma > 0$ we have*

$$\mathbb{P}[U_p(t+1) \leq \mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] - \gamma U_p(t)] \leq \exp(-\gamma^2 U_p(t)/8).$$

Proof. Note that $U_p(t+1)$ can be expressed as a function of the independent random moves of the unhappy particles. Changing the outcome of any single move changes the number of unhappy particles by at most 2. Therefore by the method of bounded differences we have

$$\mathbb{P}[U_p(t+1) \leq \mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] - \gamma U_p(t)] \leq \exp(-\gamma^2 U_p(t)/8),$$

and the proof is complete. ◀

Finally, with Lemmas 11, 12 and 13 at hand we are able to show that the duration of the dispersion process is at least exponential. The idea behind the proof is the following. We use Lemma 11 to ensure that once $U_p(t)$ dropped below $\delta n/8$ it is still $\Omega(n)$. Then, using the positive drift from Lemma 12, the number of unhappy particles increases up to $\delta n/8$. We repeat this as long as the concentration from Lemma 13 allows us to.

Proof of Theorem 2 2). Suppose $\mathcal{P}(t)$ is such that $L_t \leq n/\log n$ and $U_p(t) > cn$ for $c = \delta/16000$.

If $U_p(t) \geq \delta n/8$ then by Lemma 11 we have $\mathbb{E}[U_p(t+1)] \geq 2cn$. Together with Lemma 13 this implies

$$\mathbb{P}[U_p(t+1) < cn \mid \mathcal{P}(t)] \leq \mathbb{P}[U_p(t+1) \leq \mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] - cn \mid \mathcal{P}(t)] \leq \exp(-c^2 n/8).$$

On the other hand if $cn \leq U_p(t) < \delta n/8$ then by Lemma 12 and Lemma 13 we have

$$\begin{aligned}\mathbb{P}[U_p(t+1) < cn \mid \mathcal{P}(t)] &\leq \mathbb{P}[U_p(t+1) \leq \mathbb{E}[U_p(t+1) \mid \mathcal{P}(t)] - U_p(t)\delta/3 \mid \mathcal{P}(t)] \\ &\leq \exp(-c\delta^2 n/72).\end{aligned}$$

Together with Lemma 10 these imply that there exists a constant $\gamma > 0$ such that conditional on $\mathcal{P}(t)$ satisfying $L_t \leq n/\log n$ and $U_p(t) > cn$ we have with probability $1 - e^{-\gamma n}$ that $\mathcal{P}(t+1)$ will also satisfy these conditions. The result follows from the union bound. ◀

References

1 Colin Cooper, Andrew McDowell, Tomasz Radzik, Nicolás Rivera, and Takeharu Shiraga. Dispersion processes. *Random Structures & Algorithms*, 53(4):561–585, 2018. doi:10.1002/rsa.20822.

2 Umberto De Ambroggio, Tamás Makai, and Konstantinos Panagiotou. Dispersion on the complete graph. *Random Structures & Algorithms*, 66(1):e21266, 2025. doi:10.1002/RSA.21266.

3 Umberto De Ambroggio, Tamás Makai, Konstantinos Panagiotou, and Annika Steibel. Limit laws for critical dispersion on complete graphs, 2024. doi:10.48550/arXiv.2403.05372.

4 Alan Frieze and Wesley Pegden. A note on dispersing particles on a line. *Random Structures & Algorithms*, 53(4):586–591, 2018. doi:10.1002/RSA.20821.

5 Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. John Wiley & Sons, 2011.

6 Johannes Lengler. Drift analysis. In *Theory of evolutionary computation: Recent developments in discrete optimization*, pages 89–131. Springer, 2019. doi:10.1007/978-3-030-29414-4_2.

7 Yilun Shang. Longest distance of a non-uniform dispersion process on the infinite line. *Information Processing Letters*, 164, 2020. doi:10.1016/J.IPL.2020.106008.