A note on dispersing particles on a line

Alan Frieze*and Wesley Pegden†
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA 15213

November 6, 2017

Abstract

We consider a synchronous dispersion process introduced in [1] and we show that on the infinite line the final set of occupied sites takes up $O(n)$ space, where n is the number of particles involved.

1 Introduction

This note concerns a synchronous dispersion process introduced by Cooper, McDowell, Radzik, Rivera and Shiraga [1]. In their model, configurations of particles on the vertices of a graph evolve in discrete time; at each time step, each particle at a vertex with at least 2 particles in total moves to a neighbour chosen independently and uniformly at random. (The dispersion process thus ends at the first step when each vertex has at most one particle.) In [1], they study the behavior of this process on various graphs when begun from a configuration consisting of n particles at one vertex of the graph, with all other vertices initially empty.

They studied this process on a variety of graphs, one of which was the two-way infinite path L, with vertex set $\mathbb{Z}$ and edge set $\{\{i, i+1\}, i \in \mathbb{Z}\}$. They proved that if the initial configuration of particles consists just of n chips at the origin 0, then w.h.p. the furthest particle from the origin is at distance $O(n \log n)$ when the process stops. In this note we reduce this to $O(n)$.

Theorem 1.1. *Suppose that we begin the dispersion process on L with n particles at the origin. Then w.h.p. the furthest particle from the origin is at distance $O(n)$ when the process stops.*

*Research supported in part by NSF grant DMS1362785
†Research supported in part by NSF grant DMS1363136

2 Proof of Theorem 1.1

For our proof, we will study another equivalent process; which we call the *ordered* disperson process on L. In the ordered dispersion process, each particle has an assigned label from 1 to n; we let $X_{i,t}$ denote the position of particle i at time t. To compute one time step of the ordered process, the rules for the original dispersion process can be applied, and then the particles simply relabeled so that $i < j$ implies $X_{i,t} \leq X_{j,t}$.

Of course, it is also possible to characterize the ordered process in more complicated way without relabeling; in this view, the probability that a particle moves left or right is rarely $\frac{1}{2}$, and in general, depends both on the number of other particles in its stack, as well as the number of particles occupying the vertices adjacent to its vertex. Our proof works by analyzing this more direct (and more complicated) view of the ordered process.

The attractive feature of the ordered process is that we can define the random gaps $g_{j,t} = X_{j+1,t} - X_{j,t}$ and associated changes $\delta_{j,t} = g_{j,t+1} - g_{j,t}$. Observe that $X_{n,t} - X_{1,t} = \sum_{j=1}^{n-1} g_{j,t}$, and thus that our goal is to show that $\sum_{j=1}^{n-1} g_{j,t} = O(n)$ throughout the process.

Let $s_{j,t}$ denote the size of the stack containing particle $j \in [n]$. Define

$$I_t = \{j : g_{j,t} \geq 2\} \text{ and } J_t = \{j : s_{j,t} = 1\}.$$

Our goal is to generate random variables $h_{j,t}$ such that $g_{j,t} \leq h_{j,t}$, and such that the random variables $h_{j,t}$ are sufficiently independent for our subsequent analysis. We will first define analagous bounding random variables $\widehat{\delta}_{j,t}$ for which $\delta_{j,t} \leq \widehat{\delta}_{j,t}$. Given the present configuration, we generate independent binomials $b_{j,t} = \mathrm{Bin}(s_{j,t}, \frac{1}{2})$ for each $j \in \mathbb{Z}$. These represent the number of partices that will jump from j to $j+1$ in the unordered process. (We will only use $b_{j,t}$ when $s_{j,t} \geq 2$.)

With $s_j = s_{j,t}$, $s_{j+1} = s_{j+1,t}$, $b_j = b_{j,t}$, $b_{j+1} = b_{j+1,t}$, we define

$$\widehat{\delta}_{j,t} = \begin{cases} 2 & s_j \geq 2, s_{j+1} \geq 2, \quad b_j = 0, b_{j+1} = s_{j+1} \quad \text{(Probability } \leq 1/16.) \\ -2 & s_j \geq 2, s_{j+1} \geq 2, \quad b_j > 0, b_{j+1} < s_{j+1} \quad \text{(Probability } \geq 9/16.). \\ \\ 1 & s_j \geq 2, s_{j+1} = 1, \quad b_j = 0 \quad \text{(Probability } \leq 1/4.) \\ -1 & s_j \geq 2, s_{j+1} = 1, \quad b_j > 0 \quad \text{(Probability } \geq 3/4.) \\ \\ 1 & s_j = 1, s_{j+1} \geq 2, \quad b_{j+1} = s_{j+1} \quad \text{(Probability } \leq 1/4.) \\ -1 & s_j = 1, s_{j+1} \geq 2, \quad b_{j+1} < s_{j+1} \quad \text{(Probability } \geq 3/4.) \\ \\ 0 & \textit{otherwise.} \end{cases}$$

The probability bounds given on the right are conditioned on the configuration at time t, and thus come just from the binomials b_j, b_{j+1}.
Claim: we have $\delta_{j,t} \leq \widehat{\delta}_{j,t}$.

Proof. The gap between j and $j+1$ will increase only if all particles in j's stack move to the left and all particles in $(j+1)$'s stack move to the right (for the case where $s_j, s_{j+1} = 2$)

or if the j stack moves right and the $j+1$ stack has size 1, or the j stack has size 1 and the $j+1$ stack moves right. In the former cases, the gap increases by at most 2, and in the latter, by exactly 1. Note that we can have $\widehat{\delta}_{j,t} < \delta_{j,t}$. For example, if $b_1 = 0, b_2 = s_{j+1}$ and there is a stack S at $X_{j+1,t} + 1$ of size at least two, then a particle of S moving left, will prevent particle $j+1$ moving right in the ordered process. There are similar claims for the other positive increases in $\widehat{\delta}_{j,t}$, but not for any of the decreases. $\square$

Observe that, conditioning on the entire configuration at time t, we have

$$|\widehat{\delta}_{j,t}| \leq 2 \text{ and } \mathbf{E}(\widehat{\delta}_{j,t}) \leq -\frac{1}{2}. \tag{1}$$

Now we define an upper bound $\widehat{g}_{j,t} \geq g_{j,t}$ by setting $\widehat{g}_{j,0} = g_{j,0} = 0$ for all j, and then setting

$$\widehat{g}_{j,t+1} = \begin{cases} \max\left\{3, \widehat{g}_{j,t} + \widehat{\delta}_{j,t}\right\} & j \in I_t. \\ 3 & j \notin I_t. \end{cases}$$

Note that when $j \notin I_t$ we have $g_{j,t} < 3 \leq \widehat{g}_{j,t}$, and so $g_{j,t} \leq \widehat{g}_{j,t}$ for all j, t, with this definition.

In particular, we can now write

$$X_{n,t} - X_{1,t} = \sum_{j=1}^{n-1} g_{j,t} \leq \sum_{j=1}^{n-1} \widehat{g}_{j,t}$$

Now we can model $\widehat{g}_{j,t}$ via a random walk that begins at 3, that has a barrier at 3 and either stays where it is ($j \notin I_t$) or sometimes ($j \notin J_t$) makes a move bounded by the random variable $\widehat{\delta}_{j,t}$ which satisfies (1). It follows that there exists $0 < \rho < 1$ such that at time t we have

$$\mathbf{Pr}(\widehat{g}_{j,t} \geq 3 + k) \leq \rho^k. \tag{2}$$

(Here and elsewhere, the probability is of the event conditioned on the entire configuration at time t.)

To bound the sum $\sum_{j=1}^{n-1} \widehat{g}_{j,t}$ we use the following inequality similar to one derived in [2].

Lemma 2.1. *Let $Y_1, Y_2, \ldots, Y_m$ be independent non-negative integer random variables. Suppose that for $r \geq 1$ we have $\mathbf{Pr}(Y_r \geq k) \leq C\rho^k$, where $\rho < 1$. Let $\mu = C/(1-\rho)$. Then, if $Y = Y_1 + Y_2 + \cdots + Y_m$ then,*

$$\mathbf{Pr}(|Y - \mu m| \geq \varepsilon \mu m) \leq e^{-B\varepsilon^2 m}$$

for $0 \leq \varepsilon \leq 1$ and some $B = B(C, \rho)$.

Proof. Observe that

$$\mathbf{E}(Y_r) \leq C \sum_{k \geq 0} \rho^k = \mu, \text{ for } r \geq 1.$$

So, let $Z_r = Y_r/\mu$, so that $\mathbf{E}(Z_r) \leq 1$. Then, for $r \geq 1$ and $\lambda > 0$ such that $e^\lambda < 1/\rho$,

$$\mathbf{E}(Z_r^2 e^{\lambda Z_r}) = \sum_{k=0}^{\infty} k^2 e^{\lambda k} \, \mathbf{Pr}(Z_r = k) \leq C \sum_{k=0}^{\infty} k^2 (\rho e^\lambda)^k = \frac{C(1 - 2(\rho e^\lambda) + 3(\rho e^\lambda)^2)}{(1-\rho e^\lambda)^3} \leq \frac{2C}{(1-\rho e^\lambda)^3}.$$

Now $e^x \leq 1 + x + x^2 e^x$ and so, using the above, we have

$$\mathbf{E}(e^{\lambda Z_r}) \leq 1 + \lambda + \lambda^2 \left(1 + \frac{2C}{(1-\rho e^\lambda)^3}\right).$$

So,

$$\begin{aligned}
\mathbf{Pr}(Y/\mu \geq m + \varepsilon m) &\leq e^{-\lambda(1+\varepsilon)m} \prod_{r=1}^{m} \mathbf{E}(e^{\lambda Z_r}) \\
&\leq e^{-\lambda(1+\varepsilon)m} \exp\left\{ \left(\lambda + \lambda^2 \left(1 + \frac{2C}{(1-\rho e^\lambda)^3} \right) \right) m \right\} \\
&\leq e^{-\lambda \varepsilon m} \exp\left\{ \lambda^2 (1 + 2C\eta^{-3}) \right\}
\end{aligned}$$

assuming that

$$e^\lambda \leq (1-\eta)/\rho. \tag{3}$$

Now choose $\lambda = \varepsilon/(1 + 2C\eta^{-3})$ and $\eta = \eta(\varepsilon)$ such that (3) holds. Then

$$\mathbf{Pr}\left(Y/\mu \geq m + \varepsilon m\right) \leq \exp\left\{ -\frac{\varepsilon^2 m}{2(1+2\eta^{-3})} \right\}.$$

To bound $\mathbf{Pr}(Z/\mu \leq m - \varepsilon m)$ we use

$$\begin{aligned}
\mathbf{Pr}(Y/\mu \leq m - \varepsilon m) &\leq e^{\lambda(1-\varepsilon)m} \mathbf{E}(e^{-\lambda Y})^m \\
&\leq e^{\lambda(1-\varepsilon)m} \exp\left\{ \left(-\lambda + \lambda^2 \left(1 + \frac{3C}{(1-\rho e^\lambda)^3} \right) \right) m \right\} \\
&\leq e^{-\lambda \varepsilon m} \exp\left\{ \lambda^2 (1 + 3C\eta^{-3}) \right\},
\end{aligned}$$

and we can proceed as before. □

Now the $\widehat{g}_{j,t}$ are not independent, but we can deal with this as follows: suppose that at time t there are particles at positions $\{p_{1,t}, p_{2,t}, \ldots, p_{m_t,t}\}$. Suppose that $S_{r,t}$ denotes the set of particles at position $p_{r,t}$. Then define $P(j,t), j \in [n]$ via the equation $j \in S_{P(j,t)}$. We claim that if $j, j' \in I_t$ and $|P(j,\tau) - P(j',\tau)| \geq 2$ for all $\tau \leq t$ then $\widehat{g}_{j,t}, \widehat{g}_{j',t}$ are independent. Independence follows from the fact that $\widehat{\delta}_{j,\tau}, \widehat{\delta}_{j',\tau}$ are determined by disjoint pairs of binomials. So we let $L = \log^2 n$ and divide $[n]$ into L sets $A_i = \{j \in [n] : j \mod L = i\}$ and let $Z_{i,t} = \sum_{j \in A_i} \widehat{g}_{j,t}$.

As remarked earlier the change in $\widehat{g}_{j,\tau}, j \in A_i \cap I_\tau, \tau \leq t$ is independent of the change in $\widehat{g}_{j',\tau}, j < j' \in A_i \cap I_\tau$ as $P(j') - P(j) \geq 2$. On the other hand, if $P(j') = P(j) + 1$ then the pile S' with j' as largest element is of size at least L. Thus the probability that all particles in S' move in the same direction is at most 2^{1-L}. Thus if $\mathcal{E}$ is the event that at some time $t \leq n^{10}$ there is a pile of size at least L such that all partcles in the pile move the same way then

$$\mathbf{Pr}(\mathcal{E}) \leq n^{10} \cdot n \cdot 2^{1-L} = o(1).$$

If we condition on $\neg\mathcal{E}$ then the change in $\widehat{g}_{j,\tau}, j \in A_i \cap I_\tau, \tau \leq t$ is independent of the change in $\widehat{g}_{j',\tau}, j < j' \in A_i \cap I_\tau$ even for $P(j') - P(j) = 1$. Furthermore, these changes are conditioned downwards. This is because some of the random choices that make h increase are now precluded. We can therefore treat each $Z_{i,t}$ as the sum of indpendent random variables that satisfy (2).

Applying Lemma 2.1 we see that if $m = n/L$ then

$$\mathbf{Pr}\left(\exists i \in [L], t \leq n^{10} :\ Z_{i,t} \geq m\left(3 + \frac{2}{1-\rho}\right)\right) \leq n^{10} L e^{-C_1 m},$$

for some constant $C_1 > 0$.

It follows that with probability $1 - e^{-\Omega(n/\log^2 n)}$ we have for fixed t that

$$\sum_{j=1}^{n} \widehat{g}_{j,t} = O(n). \tag{4}$$

Our aim now is to argue that if T is the time when the process stops then w.h.p. $T \leq n^6$ so that we can apply (4) to complete the proof of Theorem 1.1. It should be mentioned that there is a proof in [1] that $T = O(n^2 \log n)$ and so what follows is only needed to make the proof self-contained.

P1 Let d_t denote the distance to the origin of the closest particle π_0 say. We observe that if Λ_t denotes the size of this closest stack, then

$$d_{t+1} - d_t = \begin{cases} 0 & \textit{probability } 1 \text{ if } \Lambda_t = 1 \\ -1 & \textit{probability } \geq 3/4 \text{ if } \Lambda_t \geq 2 \\ +1 & \textit{probability } \leq 1/4 \text{ if } \Lambda_t \geq 2 \end{cases}.$$

It follows from this and Hoeffding's theorem that w.h.p.

$$d_t = O(\log n) \text{ for } t \leq n^6. \tag{5}$$

P2 We now focus on a particular particle. If we do not re-label so that particle j always precedes particle $j+1$ then a fixed particle will do a simple random walk, interrrupted by the times when it does not make a move. The distance of a random walk from the origin after m steps can be approximated by the normal $N(0, m^{1/2}/2)$. After n^5 iterations, at least one particle will have moved n^4 times. So, after n^5 steps there is a positive probability that the walk will have reached a distance n^2 from the origin and by repetition, this will happen w.h.p. after n^6 steps. Note that different particles may be involved in different repetitions.

P3 We next observe that

$$\mathbf{Pr}(g_{j,t} \geq \log^2 n) = O(n^{-\Omega(\log n)}) \text{ for all } j \leq n, t \leq n^6.$$

This follows from $g_{j,t} \leq \widehat{g}_{j,t}$ and (2).

It follows from P1,P2, P3 that the process must finish after n^6 steps. If not then P1,P2 implies that at some time there is a particle at distance $n^2 - O(\log n)$ from particle π_0. But then there must be a gap of size $\Omega(n)$ which contradicts P3.

Now if the process finishes before n^6 steps then we can apply (4) to argue that $H_t = O(n)$ when the process finishes. Together with (5), this completes the proof of Theorem 1.1.

3 Conclusions

We have shown that the final configuration takes up $O(n)$ space w.h.p. An experiment with $n = 1000$ shows the final configuration taking up less than $1.14n$ space and so the density of occupied sites is almost 90%. It would be of some interest to try and determine the expected final density. It would also be of interest to determine the expected number of rounds more accurately. Is the $\log n$ factor in [1] needed.

One can of course ask questions concerning this problem on grids in higher dimensions. Is there a constant lower bound on the density of occupied sites in a 2-dimensional version? Can we say anything about the final shape subtended by the occupied sites? Is it a disk?

References

[1] C. Cooper, A. McDowell, T. Radzik, N. Rivera and T. Shiraga, Dispersion Processes.

[2] A.M. Frieze and W. Pegden, Traveling in randomly embedded random graphs, Proceedings of RANDOM 2017.