Natural Distributed Algorithms

- Lecture 5 -
Ant-Inspired Density Estimation

Emanuele Natale
CNRS - UCA

CdL in Informatica
Università degli Studi di Roma
"Tor Vergata"

Ants are symbol of *Biological Distributed Algorithm*

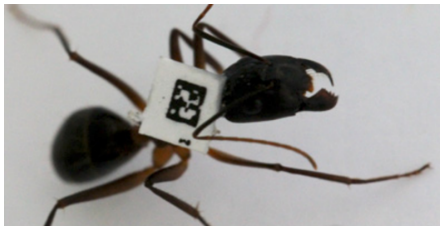
Lots of work going on:

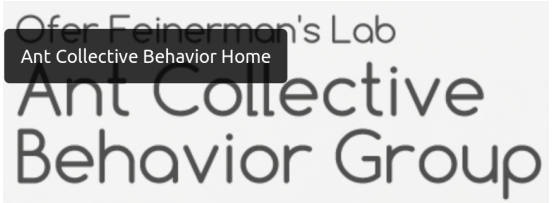
- T. Radeva, "A Symbiotic Perspective on Distributed Algorithms and Social Insects," PhD Thesis, MIT, 2017.
- A. Cornejo, A. Dornhaus, N. Lynch, and R. Nagpal, "Task Allocation in Ant Colonies," in Distributed Computing, Springer, 2014, pp. 46–60.
- Y. Afek, R. Kecher, and M. Sulamy, "Faster task allocation by idle ants," arXiv:1506.07118 [cs], Jun. 2015.
- Y. Emek, T. Langner, D. Stolz, J. Uitto, and R. Wattenhofer, "How Many Ants Does It Take to Find the Food?," in SIROCCO, Springer , 2014, pp. 263–278.
- O. Feinerman and A. Korman, "The ANTS problem," Distrib. Comput., pp. 1–20, Oct. 2016.
- L. Boczkowski, O. Feinerman, A. Korman, and E. Natale, "Limits of **Rumor Spreading** in Stochastic Populations," in ITCS, 2018.

Ants are symbol of *Biological Distributed Algorithm*

Lots of work going on:

- T. Radeva, "A Symbiotic Perspective on Distributed Algorithms and Social Insects," PhD Thesis, MIT, 2017.
- A. Cornejo, A. Dornhaus, N. Lynch, and R. Nagpal, "Task Allocation in Ant Colonies," in Distributed Computing, Springer, 2014, pp. 46–60.
- Y. Afek, R. Kecher, and M. Sulamy, "Faster task allocation by idle ants," arXiv:1506.07118 [cs], Jun. 2015.
- Y. Emek, T. Langner, D. Stolz, J. Uitto, and R. Wattenhofer, "How Many Ants Does It Take to Find the Food?," in SIROCCO, Springer , 2014, pp. 263–278.
- O. Feinerman and A. Korman, "The ANTS problem," Distrib. Comput., pp. 1–20, Oct. 2016.
- L. Boczkowski, O. Feinerman, A. Korman, and E. Natale, "Limits of **Rumor Spreading** in Stochastic Populations," in ITCS, 2018.

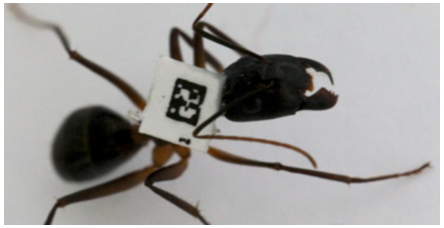




Ants are symbol of *Biological Distributed Algorithm*

Lots of work going on:

- T. Radeva, "A Symbiotic Perspective on Distributed Algorithms and Social Insects," PhD Thesis, MIT, 2017.
- A. Cornejo, A. Dornhaus, N. Lynch, and R. Nagpal, "Task Allocation in Ant Colonies," in Distributed Computing, Springer, 2014, pp. 46–60.
- Y. Afek, R. Kecher, and M. Sulamy, "Faster task allocation by idle ants," arXiv:1506.07118 [cs], Jun. 2015.
- Y. Emek, T. Langner, D. Stolz, J. Uitto, and R. Wattenhofer, "How Many Ants Does It Take to Find the Food?," in SIROCCO, Springer , 2014, pp. 263–278.
- O. Feinerman and A. Korman, "The ANTS problem," Distrib. Comput., pp. 1–20, Oct. 2016.
- L. Boczkowski, O. Feinerman, A. Korman, and E. Natale, "Limits of **Rumor Spreading** in Stochastic Populations," in ITCS, 2018.

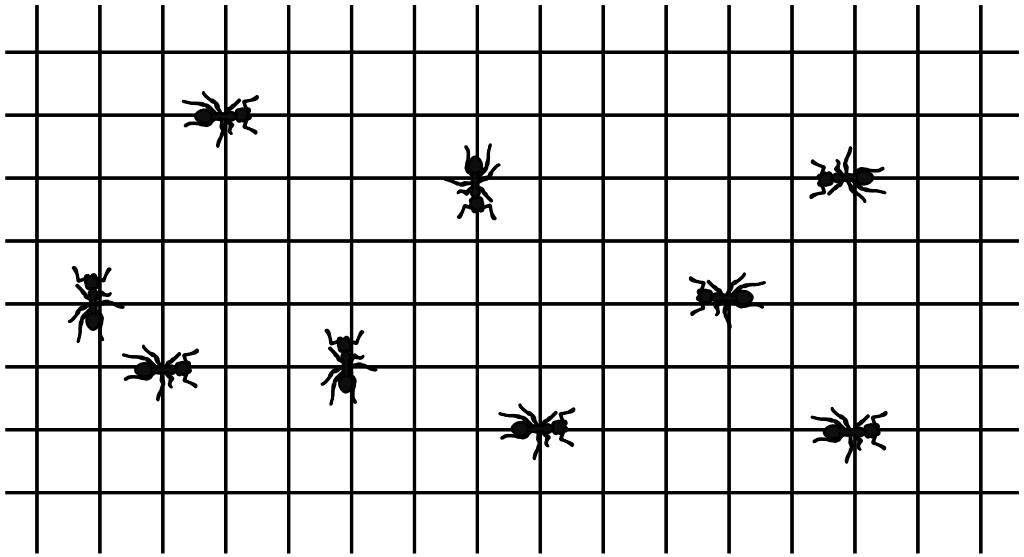


Ofer Feinerman's Lab
Ant Collective Behavior Home
Ant Collective Behavior Group

Today we talk about

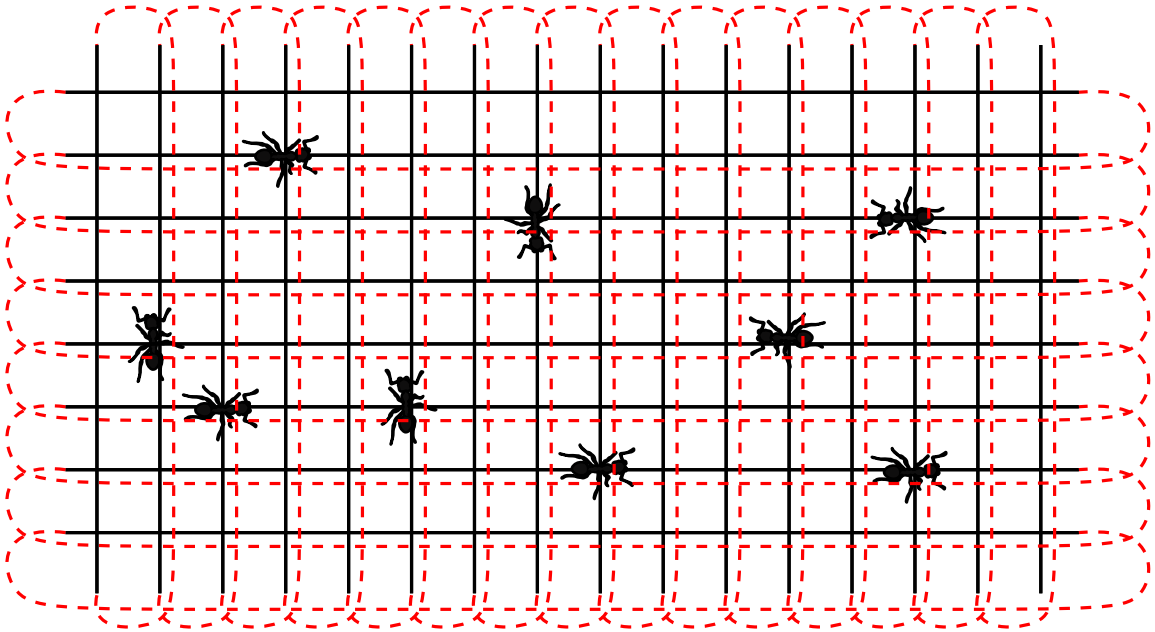
- C. Musco, H.-H. Su, and N. Lynch, "Ant-Inspired Density Estimation via Random Walks: Extended Abstract," In PODC 2016, pp. 469–478.
- C. Musco, H.-H. Su, and N. A. Lynch, "Ant-inspired density estimation via random walks," In PNAS, vol. 114, no. 40, pp. 10534–10541, Oct. 2017.

Density Estimation Problem



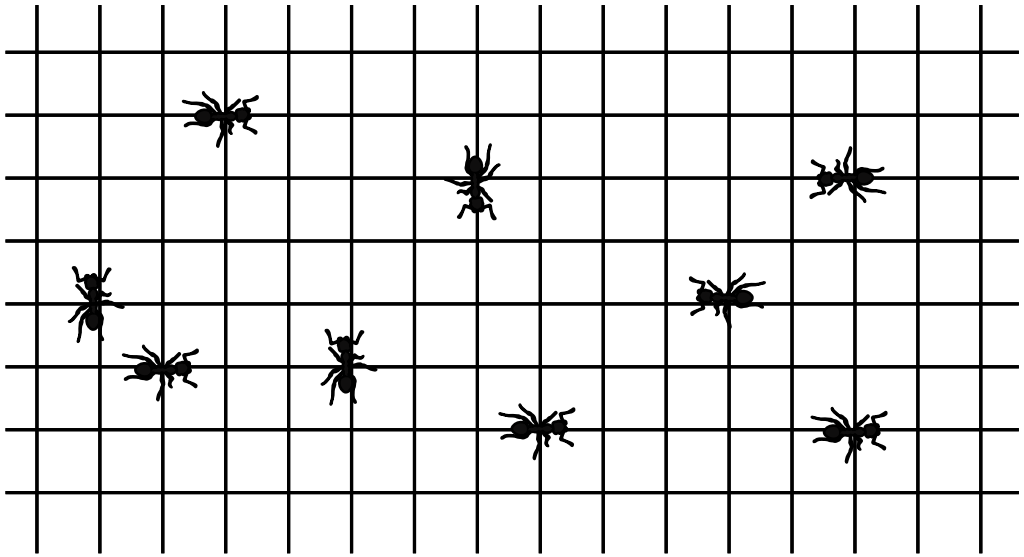
A graph (say a grid) of size $\sqrt{A} \times \sqrt{A}$, n ants. Each ant wants to learn the density $d = n/A$.

Density Estimation Problem



A graph (say a ~~grid~~ torus) of size $\sqrt{A} \times \sqrt{A}$, n ants. Each ant wants to learn the density $d = n/A$.

Density Estimation Problem

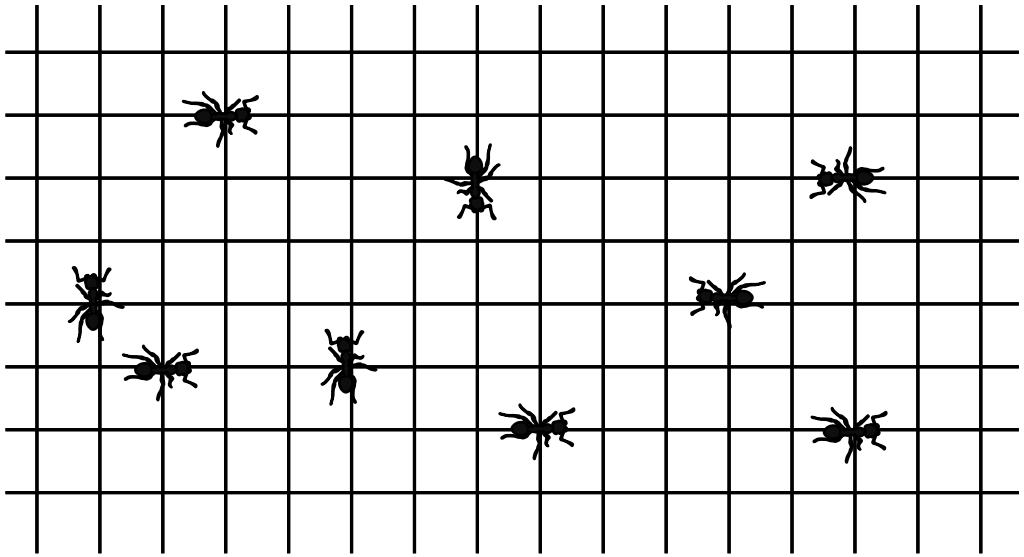


A graph (say a grid) of size $\sqrt{A} \times \sqrt{A}$, n ants. Each ant wants to learn the density $d = n/A$.

Density estimation in ants: quorum sensing during hause hunting (*temnothorax*), appraisal of enemy colony strength (*azteca*), task allocation.

How they do it?

Density Estimation Problem



A graph (say a grid) of size $\sqrt{A} \times \sqrt{A}$, n ants. Each ant wants to learn the density $d = n/A$.

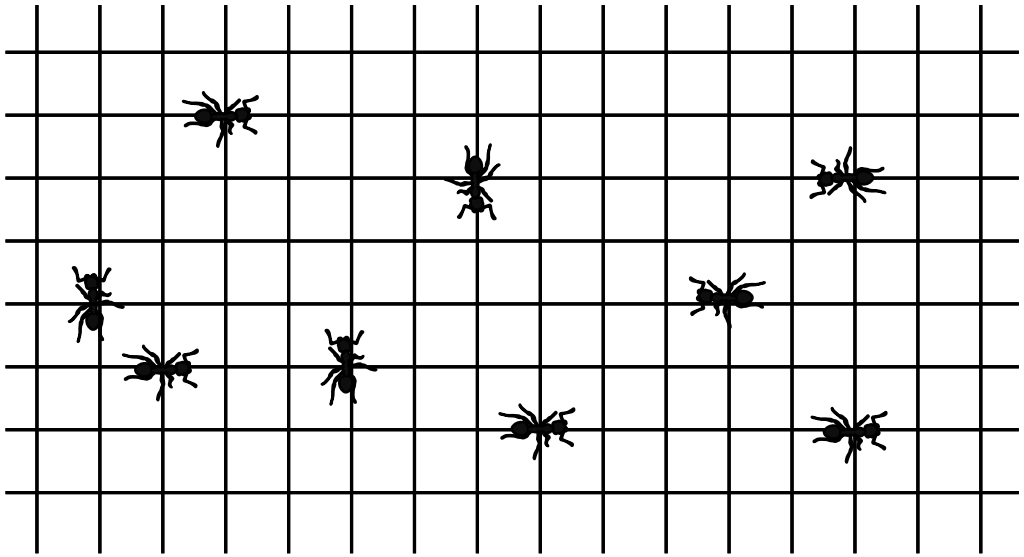
Density estimation in ants: quorum sensing during hause hunting (*temnothorax*), appraisal of enemy colony strength (*azteca*), task allocation.

How they do it?

They estimate frequency of encounters:

higher density $\Longrightarrow$ higher bumping!

Density Estimation Problem



A graph (say a grid) of size $\sqrt{A} \times \sqrt{A}$, n ants. Each ant wants to learn the density $d = n/A$.

Density estimation in ants: quorum sensing during hause hunting (*temnothorax*), appraisal of enemy colony strength (*azteca*), task allocation.

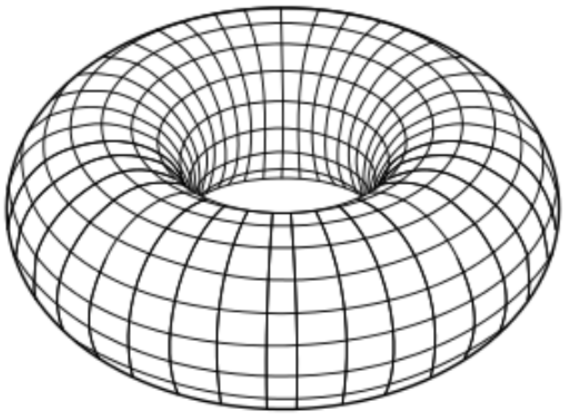
How they do it?

They estimate frequency of encounters: higher density $\Longrightarrow$ higher bumping!

Applications: Size estimation for social networks, random-walk based sampling for sensor networks, density estimation for robot swarms.

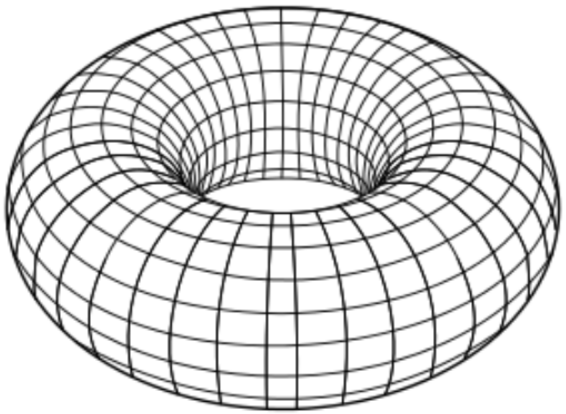
Model Definition

- Underlying graph G (2-D torus).



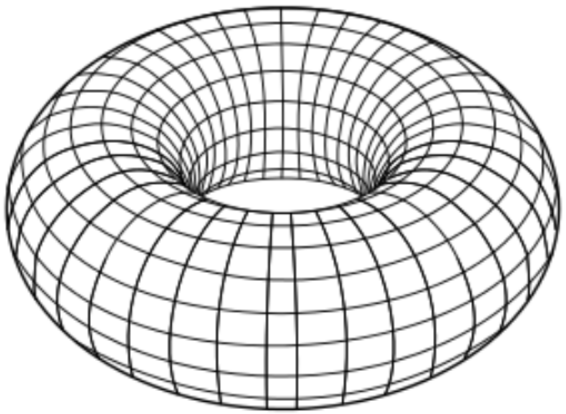
Model Definition

- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.



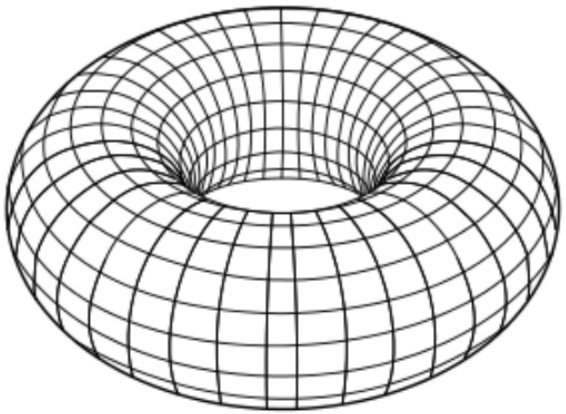
Model Definition

- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.

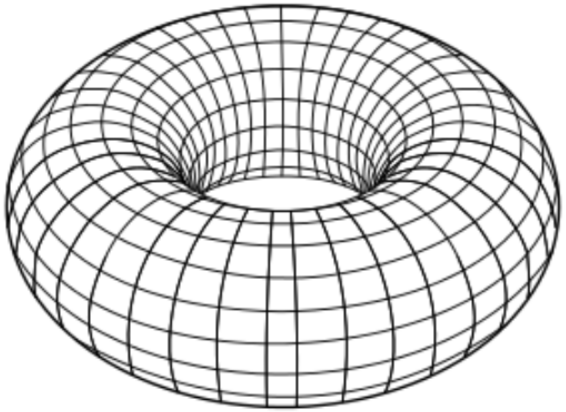


Model Definition

- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).

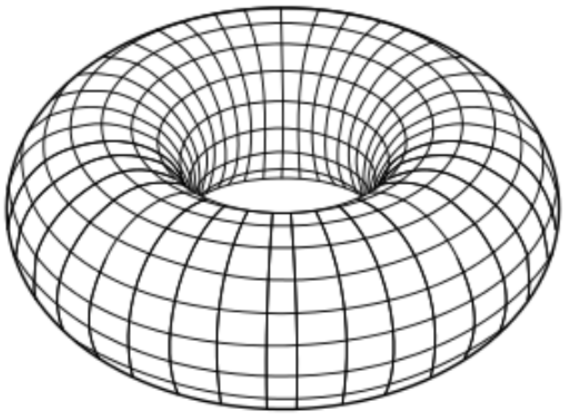


Model Definition



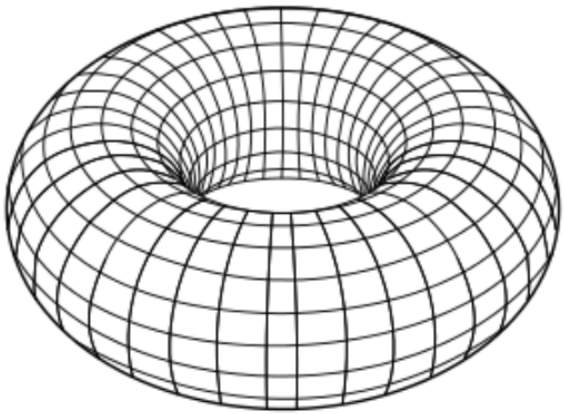
- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.

Model Definition



- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.
- The estimator of ant v after T steps is $\tilde{d} = \frac{1}{T}\sum_{t=1}^{T} c(t)$.

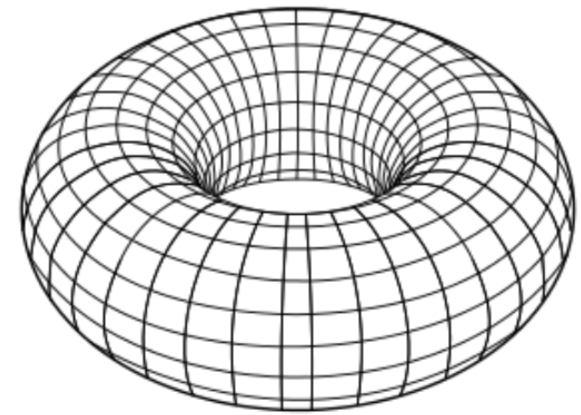
Model Definition



- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.
- The estimator of ant v after T steps is $\tilde{d} = \frac{1}{T}\sum_{t=1}^{T} c(t)$.

What is $\mathbb{E}[\tilde{d}]$?

Model Definition

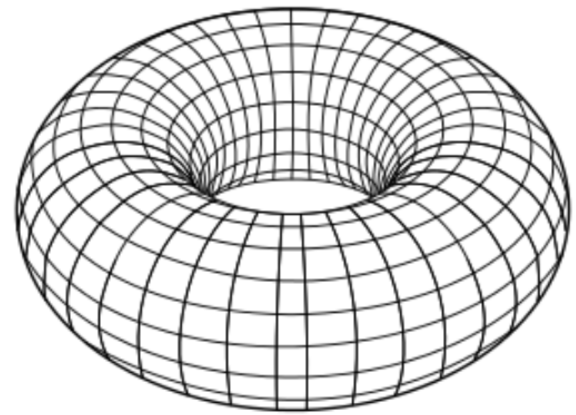


- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.
- The estimator of ant v after T steps is $\tilde{d} = \frac{1}{T}\sum_{t=1}^{T} c(t)$.

What is $\mathbb{E}[\tilde{d}]$?

Lemma. $\mathbb{E}[\tilde{d}] = d$.

Model Definition



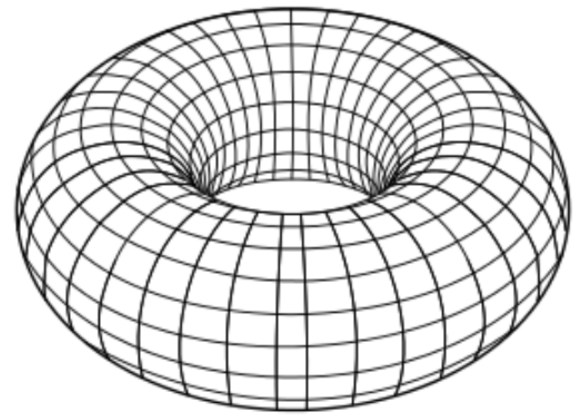
- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.
- The estimator of ant v after T steps is $\tilde{d} = \frac{1}{T}\sum_{t=1}^{T} c(t)$.

What is $\mathbb{E}[\tilde{d}]$?

Lemma. $\mathbb{E}[\tilde{d}] = d$.

Goal. If $t \geq \Theta(?)$ then $\Pr(|\tilde{d} - d| > \epsilon d) \leq \delta$.

Model Definition



- Underlying graph G (2-D torus).
- Each of the n ants is initially placed on a random node, independently from others.
- Discrete parallel time.
- At each round, each ant moves to a neighboring node chosen uniformaly at random (simple random walk).
- Fix an ant v. The only interaction of v with other ants at time t is that v can count how many ants $c(t)$ are on her node.
- The estimator of ant v after T steps is $\tilde{d} = \frac{1}{T}\sum_{t=1}^{T} c(t)$.

What is $\mathbb{E}[\tilde{d}]$?

Lemma. $\mathbb{E}[\tilde{d}] = d$.

Goal. If $t \geq \Theta(?)$ then $\Pr(|\tilde{d} - d| > \epsilon d) \leq \delta$.

The mathematical challenge: after two ants meet, they are more likely to meet again.

$c(t')$ and $c(t'')$ are not independent!

Recall on of Concentration Inequalities

A mathematician tosses n coins:
"The outcome is Binomial(n, $1/2$)."

Recall on of Concentration Inequalities

A mathematician tosses n coins:
"The outcome is Binomial(n, $1/2$)."

A computer scientist tosses n coins:
"The outcome is $\frac{n}{2} \pm \sqrt{n \log n}$ with high probability."

Markov inequality. X nonnegative r.v., then $\Pr(X \geq t) \leq \mathbb{E}[X]/t$.

Recall on of Concentration Inequalities

A mathematician tosses n coins:
"The outcome is Binomial(n, $1/2$)."

A computer scientist tosses n coins:
"The outcome is $\frac{n}{2} \pm \sqrt{n \log n}$ with high probability."

Markov inequality. X nonnegative r.v., then $\Pr(X \geq t) \leq \mathbb{E}[X]/t$.

For any non-decreasing function ψ,
$\Pr(X \geq t) = \Pr(\psi(X) \geq \psi(t)) \leq \mathbb{E}[\psi(X)]/\psi(t)$.

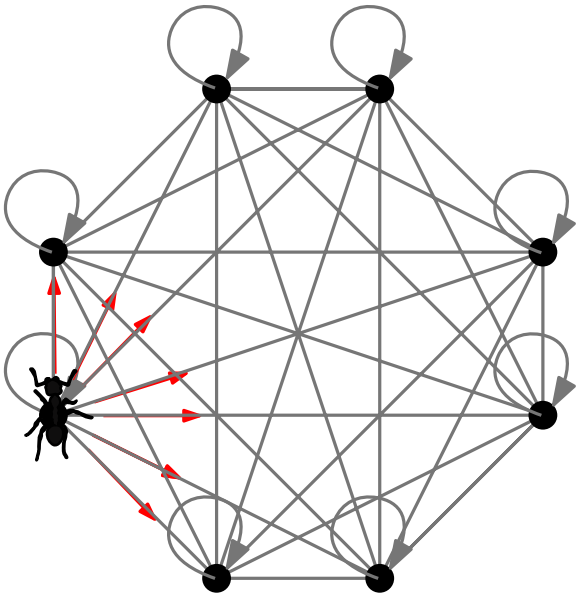
$X \leftarrow |X - \mathbb{E}X|$ and $\psi(x) = x^2 \implies$ Chebyshev inequality.

$X \leftarrow \sum_i X_i$ indip. and $\psi(X) = e^{-\lambda X} \implies$ Chernoff bounds.

Warm-up: Complete Graph

At each round each ants position is i.u.a.r.

$\implies$ $c(t')$ and $c(t'')$ are independent!

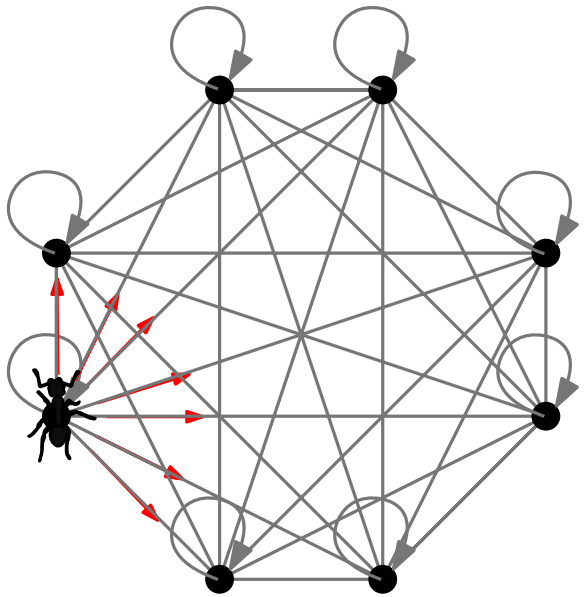


Warm-up: Complete Graph

At each round each ants position is i.u.a.r.

$\implies$ $c(t')$ and $c(t'')$ are independent!

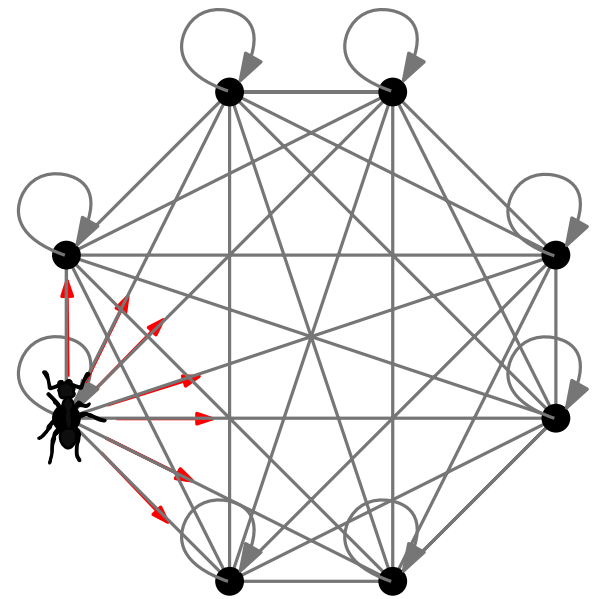
Chernoff bound. Let $X_1, \ldots, X_N$ be independent 0-1 random variables with $\Pr(X_i = 1) = p$, then for any $\epsilon \in (0, 1)$, $\Pr(|\sum_i X_i - Np| > \epsilon Np) \leq 2e^{-\frac{\epsilon^2}{3}Np}$.



Warm-up: Complete Graph

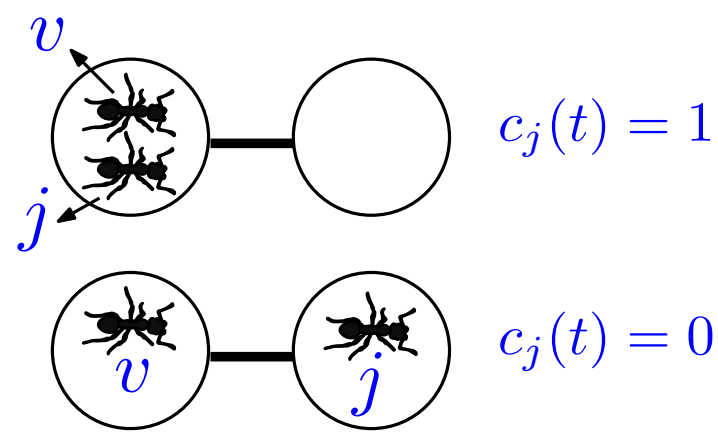
At each round each ants position is i.u.a.r.

$\implies$ $c(t')$ and $c(t'')$ are independent!



Chernoff bound. Let $X_1, \ldots, X_N$ be independent 0-1 random variables with $\Pr(X_i = 1) = p$, then for any $\epsilon \in (0, 1)$, $\Pr(|\sum_i X_i - Np| > \epsilon Np) \leq 2e^{-\frac{\epsilon^2}{3}Np}$.

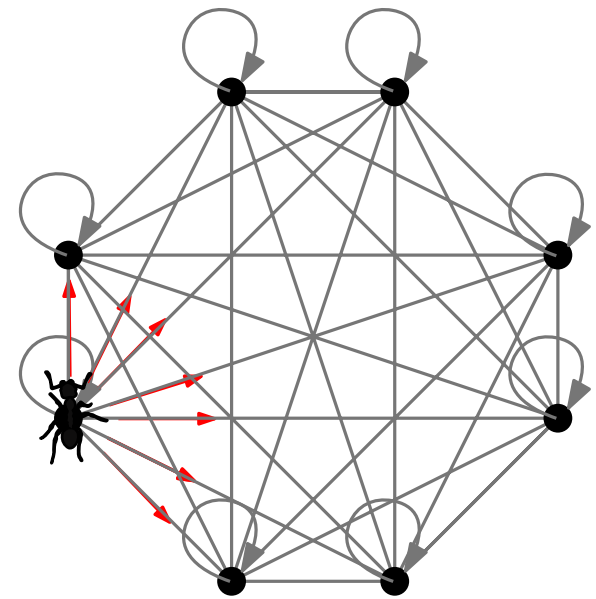
Let $c(t) = \sum_{j \neq v} c_j(t)$ where $c_j(t) = 1$ iff ant j is on v's node at time t.



Warm-up: Complete Graph

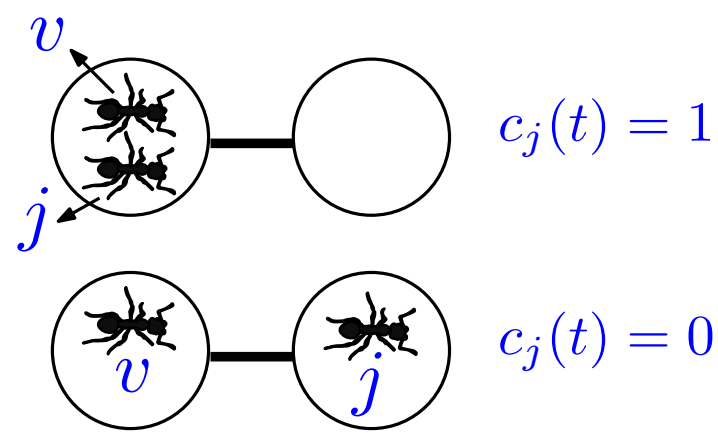
At each round each ants position is i.u.a.r.

$\implies$ $c(t')$ and $c(t'')$ are independent!



Chernoff bound. Let $X_1, \ldots, X_N$ be independent 0-1 random variables with $\Pr(X_i = 1) = p$, then for any $\epsilon \in (0,1)$, $\Pr(|\sum_i X_i - Np| > \epsilon Np) \leq 2e^{-\frac{\epsilon^2}{3}Np}$.

Let $c(t) = \sum_{j \neq v} c_j(t)$ where $c_j(t) = 1$ iff ant j is on v's node at time t.



$N = tn$, $X_{j,r} = c_j(r)$, $p = 1/A$, hence

$\Pr(|\tilde{d} - d| > \epsilon d) \leq 2e^{-\frac{\epsilon^2}{3}td} \leq \delta \implies \boxed{t = 3\log\frac{2}{\delta}/(d\epsilon^2)}$.

Main Result

Algorithm 1. Encounter Rate-Based Density Estimator

---

input: number of time steps T

$c := 0$

for $r = 1, ..., t$ **do**

$position = position + rand\{(0,1),(0,-1),(1,0),(-1,0)\}$

$c := c + count(position)$

end for

return $\tilde{d} = \frac{c}{T}$

Theorem. After running for T rounds, $T \leq A$, Algorithm 1 returns $\tilde{d}$ such that, for any $\delta > 0$, with prob $1-\delta$, $\delta d \in [(1-\epsilon)d, (1+\epsilon)d]$ for $\epsilon = \sqrt{\frac{\log \frac{1}{\delta} \log T}{Td}}$. In other words, for any $\epsilon, \delta \in (0,1)$, if $T = \Theta(\frac{\log \frac{1}{\delta} \log\log \frac{1}{\delta} \log \frac{1}{d\epsilon}}{d\epsilon^2})$, $\tilde{d}$ is a $(1 \pm \epsilon)$ multiplicative estimate of d with probability $1-\delta$.

A General Chernoff bound

General Chernoff bound (Chung-Lu). Let $X_1, ..., X_n$ be independent and $X_i \leq M$ for all i, then

$$\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \leq e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3\right)}}.$$

A General Chernoff bound

General Chernoff bound (Chung-Lu). Let $X_1, ..., X_n$ be independent and $X_i \leq M$ for all i, then

$$\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \leq e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3\right)}}.$$

Proof.

$P(\sum_i X_i - \sum_i \mathbb{E}X_i > \Delta) \leq \mathbb{E}e^{\lambda \sum_i X_i}/e^{\Delta}$.

$Ee^{\lambda \sum_i X_i} = \prod_i Ee^{\lambda X_i}$.

A General Chernoff bound

General Chernoff bound (Chung-Lu). Let $X_1, ..., X_n$ be independent and $X_i \leq M$ for all i, then

$$\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \leq e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3\right)}}.$$

Proof.

$P(\sum_i X_i - \sum_i \mathbb{E}X_i > \Delta) \leq \mathbb{E}e^{\lambda \sum_i X_i}/e^{\Delta}$.

$Ee^{\lambda \sum_i X_i} = \prod_i Ee^{\lambda X_i}$.

Let $g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}$.

It holds $g(0) = 1$, $g(y) \leq 1$ for $y < 0$ and $g(y)$ is increasing for $y \geq 0$.

A General Chernoff bound

General Chernoff bound (Chung-Lu). Let $X_1, ..., X_n$ be independent and $X_i \leq M$ for all i, then

$$\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \leq e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3\right)}}.$$

Proof.

$P(\sum_i X_i - \sum_i \mathbb{E}X_i > \Delta) \leq \mathbb{E}e^{\lambda \sum_i X_i}/e^{\Delta}$.

$Ee^{\lambda \sum_i X_i} = \prod_i Ee^{\lambda X_i}$.

Let $g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}$.

It holds $g(0) = 1$, $g(y) \leq 1$ for $y < 0$ and $g(y)$ is increasing for $y \geq 0$.

Since $k! \geq 2 \cdot 3^{k-2}$, $g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} \leq \sum_{k=2}^{\infty} \frac{y^{k-2}}{3^{k-2}} = \frac{1}{1-\frac{y}{3}}$ for $y < 3$.

A General Chernoff bound

We have

$$\begin{aligned}
\mathbb{E}\left(e^{\lambda \sum_i X}\right) &= \prod_i \mathbb{E}\left(e^{\lambda X_i}\right) = \prod_i \mathbb{E}\left(\sum_{k=0}^{\infty} \frac{\lambda^k X_i^k}{k!}\right) \\
&= \prod_i \mathbb{E}\left(1 + \lambda X_i + \frac{1}{2}\lambda^2 X_i^2 g\left(\lambda X_i\right)\right) \\
&\leq \prod_i \left(1 + \lambda \mathbb{E}\left(X_i\right) + \frac{1}{2}\lambda^2 \mathbb{E}\left(X_i^2\right) g\left(\lambda M\right)\right) \\
&\leq \prod_i e^{\lambda \mathbb{E}(X_i) + \frac{1}{2}\lambda^2 \mathbb{E}\left(X_i^2\right) g(\lambda M)} \\
&= e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)}.
\end{aligned}$$

Hence, for λ satisfying $\lambda M < 3$, we have...

A General Chernoff bound

$$
\begin{aligned}
&\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \\
&= \Pr\left(e^{\lambda X} \geq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \lambda\Delta}\right) \\
&\leq e^{-\lambda \mathbb{E}\left(\sum_i X_i\right) - \lambda\Delta} \mathbb{E}\left(e^{\lambda X}\right) \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)} \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 \frac{\sum_i \mathbb{E}\left(X_i^2\right)}{1 - \lambda M/3}}.
\end{aligned}
$$

A General Chernoff bound

$$
\begin{aligned}
&\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \\
&= \Pr\left(e^{\lambda X} \geq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \lambda\Delta}\right) \\
&\leq e^{-\lambda \mathbb{E}\left(\sum_i X_i\right) - \lambda\Delta} \mathbb{E}\left(e^{\lambda X}\right) \longrightarrow \leq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)} \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)} \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 \frac{\sum_i \mathbb{E}\left(X_i^2\right)}{1 - \lambda M/3}}.
\end{aligned}
$$

A General Chernoff bound

$$\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right)$$

$$= \Pr\left(e^{\lambda X} \geq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \lambda\Delta}\right)$$

$$\leq e^{-\lambda \mathbb{E}\left(\sum_i X_i\right) - \lambda\Delta} \mathbb{E}\left(e^{\lambda X}\right) \longrightarrow \leq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)}$$

$$\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)}$$

$$g(\lambda M) \leq \frac{1}{1 - \frac{\lambda M}{3}}$$

$$\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 \frac{\sum_i \mathbb{E}\left(X_i^2\right)}{1 - \lambda M/3}}.$$

A General Chernoff bound

$$
\begin{aligned}
&\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) \\
&= \Pr\left(e^{\lambda X} \geq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \lambda\Delta}\right) \\
&\leq e^{-\lambda \mathbb{E}\left(\sum_i X_i\right) - \lambda\Delta} \mathbb{E}\left(e^{\lambda X}\right) \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)} \\
&\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 \frac{\sum_i \mathbb{E}\left(X_i^2\right)}{1 - \lambda M/3}}.
\end{aligned}
$$

$$\leq e^{\lambda \mathbb{E}\left(\sum_i X_i\right) + \frac{1}{2}\lambda^2 g(\lambda M) \sum_i \mathbb{E}\left(X_i^2\right)}$$

$$g(\lambda M) \leq \frac{1}{1 - \frac{\lambda M}{3}}$$

Choosing $\lambda = \frac{\Delta}{\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3}$, we have $\lambda M < 3$ and

$$
\begin{aligned}
\Pr\left(\sum_i X_i \geq \mathbb{E}\left(\sum_i X_i\right) + \Delta\right) &\leq e^{-\lambda\Delta + \frac{1}{2}\lambda^2 \frac{\sum_i \mathbb{E}\left(X_i^2\right)}{1 - \lambda M/3}} \\
&\leq e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}\left(X_i^2\right) + M\Delta/3\right)}} \qquad \square
\end{aligned}
$$

Proof Ingredients of Theorem 1

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

Proof Ingredients of Theorem 1

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Proof Ingredients of Theorem 1

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Collision Moment Lemma. For $j \in [1, ..., n]$, let $\bar{c}_j \stackrel{def}{=} c_j - \mathbb{E}[c_j]$. For all $k \geq 2$, assuming $t \leq A$, $|\mathbb{E}[\bar{c}_j^k]| = \mathcal{O}(\frac{t}{A} k! \log^{k-1} t)$.

Proof Ingredients of Theorem 1

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Collision Moment Lemma. For $j \in [1, ..., n]$, let $\bar{c}_j \stackrel{def}{=} c_j - \mathbb{E}[c_j]$. For all $k \geq 2$, assuming $t \leq A$, $|\mathbb{E}[\bar{c}_j^k]| = \mathcal{O}(\frac{t}{A} k! \log^{k-1} t)$.

Bernstein Inequality. If $|E[\bar{c}_j^k]| \leq \frac{1}{2} k! \sigma^2 b^{k-2}$ for each $k \geq 2$, then

$$\Pr(\sum_i X_i - \sum_i \mathbb{E} X_i \geq t) \leq e^{-\frac{t^2}{2(\sigma^2 + bt)}}.$$

Proof Ingredients of Theorem 1

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Collision Moment Lemma. For $j \in [1, ..., n]$, let $\bar{c}_j \stackrel{def}{=} c_j - \mathbb{E}[c_j]$. For all $k \geq 2$, assuming $t \leq A$, $|\mathbb{E}[\bar{c}_j^k]| = \mathcal{O}(\frac{t}{A} k! \log^{k-1} t)$.

Bernstein Inequality. If $|E[\bar{c}_j^k]| \leq \frac{1}{2} k! \sigma^2 b^{k-2}$ for each $k \geq 2$, then

$$\Pr(\sum_i X_i - \sum_i \mathbb{E} X_i \geq t) \leq e^{-\frac{t^2}{2(\sigma^2 + bt)}}.$$

Remark. Proofs can be revisited to estimate probability that single random walk return on a given node (equalization).

Proof of Re-collision Lemma

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

Two random walkers, a_1 and a_2.
Let M_x and M_y be the steps on x and y direction ($M_x + M_y = 2m$).
Let $\mathcal{C} =$"they re-collide after t steps", and $\mathcal{C}_x$, and $\mathcal{C}_y$, the event that they end with same x, and y.

Proof of Re-collision Lemma

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

Two random walkers, a_1 and a_2.
Let M_x and M_y be the steps on x and y direction ($M_x + M_y = 2m$).
Let $\mathcal{C} =$"they re-collide after t steps", and $\mathcal{C}_x$, and $\mathcal{C}_y$, the event that they end with same x, and y.

$\Pr(\mathcal{C} \mid M_x = m_x, M_y = m_y) = \Pr(\mathcal{C}_x \mid M_x = m_x) \Pr(\mathcal{C}_y \mid M_y = m_y)$.

Proof of Re-collision Lemma

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

Two random walkers, a_1 and a_2.
Let M_x and M_y be the steps on x and y direction ($M_x + M_y = 2m$).
Let $\mathcal{C} =$"they re-collide after t steps", and $\mathcal{C}_x$, and $\mathcal{C}_y$, the event that they end with same x, and y.

$\Pr(\mathcal{C} \mid M_x = m_x, M_y = m_y) = \Pr(\mathcal{C}_x \mid M_x = m_x) \Pr(\mathcal{C}_y \mid M_y = m_y)$.

Wlog, we look at $\mathcal{C}_x$.
Let $\mathcal{C}_x^1$ and $\mathcal{C}_x^2$ be the events "same x without displacement" and "same x with displacement" (displacement=wrapping around torus), so
$\Pr(\mathcal{C}_x \mid M_x = m_x) = \Pr(\mathcal{C}_x^1 \mid M_x = m_x) + \Pr(\mathcal{C}_x^2 \mid M_x = m_x)$.

Proof of Re-collision Lemma

Re-collision Lemma. Consider two agents a_1 and a_2 randomly walking on a $\sqrt{A} \times \sqrt{A}$ torus. If a_1 and a_2 collide at time t, the prob. that they collide again in round $m + t$ is $\mathcal{O}(\frac{1}{m+1}) + \mathcal{O}(\frac{1}{A})$.

Two random walkers, a_1 and a_2.
Let M_x and M_y be the steps on x and y direction ($M_x + M_y = 2m$).
Let $\mathcal{C} =$"they re-collide after t steps", and $\mathcal{C}_x$, and $\mathcal{C}_y$, the event that they end with same x, and y.

$\Pr(\mathcal{C} \mid M_x = m_x, M_y = m_y) = \Pr(\mathcal{C}_x \mid M_x = m_x) \Pr(\mathcal{C}_y \mid M_y = m_y)$.

Wlog, we look at $\mathcal{C}_x$.
Let $\mathcal{C}_x^1$ and $\mathcal{C}_x^2$ be the events "same x without displacement" and "same x with displacement" (displacement=wrapping around torus), so
$\Pr(\mathcal{C}_x \mid M_x = m_x) = \Pr(\mathcal{C}_x^1 \mid M_x = m_x) + \Pr(\mathcal{C}_x^2 \mid M_x = m_x)$.

The first summand means that the random walk comes back to the origin: $\Pr(\mathcal{C}_x^1 \mid M_x = m_x) = \binom{m_x}{m_x/2}(\frac{1}{2})^{m_x} = \frac{m_x!}{((m_x/2)!)^2}(\frac{1}{2})^{m_x}$.

Proof of Re-collision Lemma

Assuming m_x even and using Stirling $n! = \sqrt{2\pi n}(\frac{n}{e})^n(1 + \mathcal{O}(\frac{1}{n}))$, we get $\Pr(\mathcal{C}_x^1 \mid M_x = m_x) = \Theta(1/\sqrt{m_x + 1})$.

Proof of Re-collision Lemma

Assuming m_x even and using Stirling $n! = \sqrt{2\pi n}(\frac{n}{e})^n(1+\mathcal{O}(\frac{1}{n}))$, we get $\Pr(\mathcal{C}_x^1 \mid M_x = m_x) = \Theta(1/\sqrt{m_x+1})$.

As for C_x^2, $\Pr(\mathcal{C}_x^2 \mid M_x = m_x) = 2(\frac{1}{2})^{m_x} \sum_{c=1}^{\lfloor \frac{m_x}{\sqrt{A}} \rfloor} \binom{m_x}{(m_x - c\sqrt{A})/2}$.

Proof of Re-collision Lemma

Assuming m_x even and using Stirling $n! = \sqrt{2\pi n}(\frac{n}{e})^n(1+\mathcal{O}(\frac{1}{n}))$, we get $\Pr(\mathcal{C}_x^1 \mid M_x = m_x) = \Theta(1/\sqrt{m_x+1})$.

As for C_x^2, $\Pr(\mathcal{C}_x^2 \mid M_x = m_x) = 2(\frac{1}{2})^{m_x} \sum_{c=1}^{\lfloor \frac{m_x}{\sqrt{A}} \rfloor} \binom{m_x}{(m_x - c\sqrt{A})/2}$.

For $i \in [1, ..., \sqrt{A}-1]$,
let $\mathcal{D}_x^i =$"the walk is i steps clockwise from start after m_x steps". It holds

$$\Pr[\mathcal{D}_x^i | M_x = m_x] = \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{c=-\left\lfloor \frac{m_x+i}{\sqrt{A}} \right\rfloor}^{\left\lfloor \frac{m_x-i}{\sqrt{A}} \right\rfloor} \binom{m_x}{\frac{m_x+i+c\sqrt{A}}{2}}$$

$$\geq \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{c=-\left\lfloor \frac{m_x+i}{\sqrt{A}} \right\rfloor}^{-1} \binom{m_x}{\frac{m_x+i+c\sqrt{A}}{2}} = \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{c=1}^{\left\lfloor \frac{m_x}{\sqrt{A}} \right\rfloor} \binom{m_x}{\frac{m_x+i-c\sqrt{A}}{2}}.$$

Proof of Re-collision Lemma

For any $i \in [1, ..., \sqrt{A} - 1]$, and any $c \geq 1$, $\frac{m_x + i - c\sqrt{A}}{2}$ is closer to $\frac{m_x}{2}$ than $\frac{m_x - c\sqrt{A}}{2}$ is, so

$$\binom{m_x}{\frac{m_x+i-c\sqrt{A}}{2}} > \binom{m_x}{\frac{m_x-c\sqrt{A}}{2}}$$

as long as $\frac{m_x+i-c\sqrt{A}}{2}$ is an integer. This allows us to lower bound $\Pr[\mathcal{D}_x^i | M_x = m_x]$ using $\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right]$. Let $\mathcal{E}_{i,c}$ equal 1 if $\frac{m_x+i-c\sqrt{A}}{2}$ is an integer and 0 otherwise. Since $\mathcal{C}_x^2$ and each $\mathcal{D}_x^i$ are disjoint events:

$$\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] + \sum_{i=1}^{\sqrt{A}-1} \Pr\left[\mathcal{D}_x^i | M_x = m_x\right] \leq 1$$

$$\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] + \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{i=1}^{\sqrt{A}-1} \left(\sum_{c=1}^{\left\lfloor \frac{m_x}{\sqrt{A}} \right\rfloor} \binom{m_x}{\frac{m_x+i-c\sqrt{A}}{2}} \right) \leq 1$$

Proof of Re-collision Lemma

$$\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] + \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{i=1}^{\sqrt{A}-1} \left(\sum_{c=1}^{\left\lfloor \frac{m_x}{\sqrt{A}} \right\rfloor} \binom{m_x}{\frac{m_x+i-c\sqrt{A}}{2}} \right) \leq 1$$

$$\Downarrow$$

$$\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] + \left(\frac{1}{2}\right)^{m_x} \cdot \sum_{c=1}^{\left\lfloor \frac{m_x}{\sqrt{A}} \right\rfloor} \left(\binom{m_x}{\frac{m_x-c\sqrt{A}}{2}} \cdot \sum_{i=1}^{\sqrt{A}-1} \mathcal{E}_{i,c} \right) \leq 1$$

$$\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] \cdot \Theta(\sqrt{A}) \leq 1.$$

The last step follows from combining the last with the fact that $\sum_{i=1}^{\sqrt{A}-1} \mathcal{E}_{i,c} = \Theta\left(\sqrt{A}\right)$ for all c since $\frac{m_x+i-c\sqrt{A}}{2}$ is integral for half the possible $i \in [1, ..., \sqrt{A}-1]$. Rearranging, we have $\Pr\left[\mathcal{C}_x^2 | M_x = m_x\right] = O\left(\frac{1}{\sqrt{A}}\right)$.

Proof of Re-collision Lemma

Combining our bounds for $\mathcal{C}_x^1$ and $\mathcal{C}_x^2$,
$\Pr\left[\mathcal{C}_x | M_x = m_x\right] = \Theta\left(\frac{1}{\sqrt{m_x+1}}\right) + O\left(\frac{1}{\sqrt{A}}\right).$

Proof of Re-collision Lemma

Combining our bounds for $\mathcal{C}_x^1$ and $\mathcal{C}_x^2$,
$\Pr\left[\mathcal{C}_x | M_x = m_x\right] = \Theta\left(\frac{1}{\sqrt{m_x+1}}\right) + O\left(\frac{1}{\sqrt{A}}\right).$

Identical bounds hold for the y direction and by saparating horizantal/vertical axis we have:

$$\Pr\left[\mathcal{C} | M_x = m_x, M_y = m_y\right] = \Theta\left(\frac{1}{\sqrt{(m_x+1)(m_y+1)}}\right) + O\left(\frac{1}{\sqrt{A(m_x+1)}} + \frac{1}{\sqrt{A(m_y+1)}}\right) + O\left(\frac{1}{A}\right).$$

Proof of Re-collision Lemma

Our final step is to remove the conditioning on M_x and M_y. Since direction is chosen independently and uniformly at random for each step, $\mathbf{E}\left[M\right]_x = \mathbf{E}\left[M\right]_y = m$. By a standard Chernoff bound:

$$\Pr[M_x \leq m/2] \leq 2e^{-(1/2)^2 \cdot m/2} = O\left(\frac{1}{m+1}\right).$$

(using $m+1$ instead of m to cover the $m = 0$ case).

An identical bound holds for M_y, and so, except with probability $O\left(\frac{1}{m+1}\right)$ both are $\geq m/2$. We get:

$$\begin{aligned}\Pr\left[\mathcal{C}\right] &= \Theta\left(\frac{1}{m+1}\right) + O\left(\frac{1}{\sqrt{A(m+1)}}\right) + O\left(\frac{1}{A}\right) \\ &= \Theta\left(\frac{1}{m+1}\right) + O\left(\frac{1}{A}\right). \qquad \square\end{aligned}$$

Proof of First Collision Lemma

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Using the fact that c_j is identically distributed for all j,

$$\mathbb{E}[\tilde{d}] = d = \frac{1}{t} \cdot \mathbb{E}[\sum_{i=1}^{n} c_i] = \frac{n}{t} \cdot \mathbb{E}[c_j] = \frac{n}{t} \cdot \Pr[c_j \geq 1] \cdot \mathbb{E}[c_j | c_j \geq 1],$$

that is

$$\frac{n}{A} = d = \frac{n}{t} \cdot \Pr[c_j \geq 1] \cdot \mathbb{E}[c_j | c_j \geq 1].$$

Proof of First Collision Lemma

First-collision Lemma. Assuming $t \leq A$, for all $j \in [1, ..., n]$, $\Pr[c_j \geq 1] = \Theta(\frac{t}{A \log t})$.

Using the fact that c_j is identically distributed for all j,

$$\mathbb{E}[\tilde{d}] = d = \frac{1}{t} \cdot \mathbb{E}[\sum_{i=1}^{n} c_i] = \frac{n}{t} \cdot \mathbb{E}[c_j] = \frac{n}{t} \cdot \Pr[c_j \geq 1] \cdot \mathbb{E}[c_j | c_j \geq 1],$$

that is

$$\frac{n}{A} = d = \frac{n}{t} \cdot \Pr[c_j \geq 1] \cdot \mathbb{E}[c_j | c_j \geq 1].$$

Rearranging gives:

$$\Pr[c_j \geq 1] = \frac{t}{A \cdot \mathbb{E}[c_j | c_j \geq 1]}.$$

Proof of First Collision Lemma

To compute $\mathbb{E}[c_j | c_j \geq 1]$, we use Re-collision Lemma and linearity of expectation. Since $t \leq A$, the $O\left(\frac{1}{A}\right)$ term in Re-collision Lemma is absorbed into the $\Theta\left(\frac{1}{m+1}\right)$. Let $r \leq t$ be the first round that the two agents collide. We have:

$$\mathbb{E}[c_j | c_j \geq 1] = \sum_{m=0}^{t-r} \Theta\left(\frac{1}{m+1}\right) = \Theta\left(\log(t-r)\right).$$

Proof of First Collision Lemma

To compute $\mathbb{E}[c_j | c_j \geq 1]$, we use Re-collision Lemma and linearity of expectation. Since $t \leq A$, the $O\left(\frac{1}{A}\right)$ term in Re-collision Lemma is absorbed into the $\Theta\left(\frac{1}{m+1}\right)$. Let $r \leq t$ be the first round that the two agents collide. We have:

$$\mathbb{E}[c_j | c_j \geq 1] = \sum_{m=0}^{t-r} \Theta\left(\frac{1}{m+1}\right) = \Theta\left(\log(t-r)\right).$$

- After any round the agents are located at uniformly and independently chosen positions, so collide with probability exactly $1/A$.
- The probability that the *first* collision between the agents happens in a given round can only decrease as we consider a round later in time.
- At least $1/2$ of the time that $c_j \geq 1$, there is a collision in the first $t/2$ rounds.

Thanks to the previous calculations,
$\mathbb{E}[c_j | c_j \geq 1] = \Theta\left(\log(t - t/2)\right) = \Theta\left(\log t\right)$, hence
$\Pr\left[c_j \geq 1\right] = \Theta\left(\frac{t}{A \cdot \log t}\right)$, completing the proof. $\square$

Proof of Collision Moment Lemma

Collision Moment Lemma. For $j \in [1, ..., n]$, let $\bar{c}_j \stackrel{def}{=} c_j - \mathbb{E}c_j$.
For all $k \geq 2$, assuming $t \leq A$, $\mathbb{E}[\bar{c}_j^k] = \mathcal{O}(\frac{t}{A} k! \log^{k-1} t)$.

We expand $\mathbb{E}[\bar{c}_j^k] = \Pr[c_j \geq 1] \cdot \mathbb{E}[\bar{c}_j^k | c_j \geq 1] + \Pr[c_j = 0] \cdot \mathbb{E}[\bar{c}_j^k | c_j = 0]$,
and so by First Collision Lemma:

$$\mathbb{E}\left[\bar{c}_j^k\right] = O\left(\frac{t}{A \log t} \cdot \mathbb{E}\left[\bar{c}_j^k | c_j \geq 1\right] + \mathbb{E}\left[\bar{c}_j^k | c_j = 0\right]\right).$$

Proof of Collision Moment Lemma

Collision Moment Lemma. For $j \in [1, ..., n]$, let $\bar{c}_j \stackrel{def}{=} c_j - \mathbb{E}c_j$.
For all $k \geq 2$, assuming $t \leq A$, $\mathbb{E}[\bar{c}_j^k] = \mathcal{O}(\frac{t}{A} k! \log^{k-1} t)$.

We expand $\mathbb{E}[\bar{c}_j^k] = \Pr[c_j \geq 1] \cdot \mathbb{E}[\bar{c}_j^k | c_j \geq 1] + \Pr[c_j = 0] \cdot \mathbb{E}[\bar{c}_j^k | c_j = 0]$,
and so by First Collision Lemma:

$$\mathbb{E}\left[\bar{c}_j^k\right] = O\left(\frac{t}{A \log t} \cdot \mathbb{E}\left[\bar{c}_j^k | c_j \geq 1\right] + \mathbb{E}\left[\bar{c}_j^k | c_j = 0\right]\right).$$

$\mathbb{E}\left[\bar{c}_j^k | c_j = 0\right] = (\mathbb{E}c_j)^k = (t/A)^k \leq \frac{t}{A} k! \log^{k-1} t$ for all $k \geq 2$.

Moreover $\mathbb{E}\left[\bar{c}_j^k | c_j \geq 1\right] \leq \mathbb{E}\left[c_j^k | c_j \geq 1\right]$, since $\mathbb{E}c_j = \frac{t}{A} \leq 1$.

To prove the lemma, it just remains to show that
$\mathbb{E}\left[c_j^k | c_j \geq 1\right] = O\left(k! \log^k t\right)$.

Proof of Collision Moment Lemma

Conditioning on $c_j \geq 1$, we know the agents have an initial collision in some round $t' \leq t$. We split c_j over rounds: $c_j = \sum_{r=t'}^{t} c_j(r) \leq \sum_{r=t'}^{t'+t-1} c_j(r)$. To simplify notation we relabel round t' round 1 and so round $t' + t - 1$ becomes round t. Expanding c_j^k out fully using the summation:

$$\begin{aligned}\mathbb{E}\left[c_j^k\right] &= \mathbb{E}\left[\sum_{r_1=1}^{t}\sum_{r_2=1}^{t}\ldots\sum_{r_k=1}^{t} c_j(r_1)c_j(r_2)\ldots c_j(r_k)\right] \\ &= \sum_{r_1=1}^{t}\sum_{r_2=1}^{t}\ldots\sum_{r_k=1}^{t} \mathbb{E}\left[c_j(r_1)c_j(r_2)\ldots c_j(r_k)\right].\end{aligned}$$

Proof of Collision Moment Lemma

$\mathbb{E}\left[c_{r_1}(j)c_{r_2}(j)...c_{r_k}(j)\right]$ is just the probability that the two agents collide in each of rounds $r_1, r_2, ...r_k$. Assume w.l.o.g. that $r_1 \leq r_2 \leq ... \leq r_k$. By Re-collision Lemma this is: $O\left(\frac{1}{r_1(r_2-r_1+1)(r_3-r_2+1)...(r_k-r_{k-1}+1)}\right)$. So we can rewrite, by linearity of expectation:

$$\left.\right] = k! \sum_{r_1=1}^{t} \sum_{r_2=r_1}^{t} ... \sum_{r_k=r_{k-1}}^{t} O\left(\frac{1}{r_1(r_2-r_1+1)(r_3-r_2+1)...(r_k-r_{k-1}+1)}\right).$$

Proof of Collision Moment Lemma

The $k!$ comes from the fact that in this sum we only have *ordered* k-tuples and so need to multiple by $k!$ to account for the fact that the original sum is over *unordered* k-tuples. We can bound:

$$\sum_{r_k=r_{k-1}}^{t} \frac{1}{r_k - r_{k-1} + 1} = 1 + \frac{1}{2} + ... + \frac{1}{t} = O(\log t)$$

so rearranging the sum and simplifying gives:

$$\begin{aligned}\mathbb{E}\left[c_j^k\right] &= k! \sum_{r_1=1}^{t} \frac{1}{r_1} \sum_{r_2=r_1+1}^{t} \frac{1}{r_2 - r_1} ... \sum_{r_k=r_{k-1}+1}^{t} \frac{1}{r_k - r_{k-1}} \\ &= k! \sum_{r_1=1}^{t} \frac{1}{r_1} \sum_{r_2=r_1}^{t} \frac{1}{r_2 - r_1 + 1} ... \sum_{r_{k-1}=r_{k-2}}^{t} \frac{1}{r_{k-2} - r_{k-1} + 1} \cdot O(\log t).\end{aligned}$$

Proof of Collision Moment Lemma

The $k!$ comes from the fact that in this sum we only have *ordered* k-tuples and so need to multiple by $k!$ to account for the fact that the original sum is over *unordered* k-tuples. We can bound:

$$\sum_{r_k=r_{k-1}}^{t} \frac{1}{r_k - r_{k-1} + 1} = 1 + \frac{1}{2} + \ldots + \frac{1}{t} = O(\log t)$$

so rearranging the sum and simplifying gives:

$$\begin{aligned}\mathbb{E}\left[c_j^k\right] &= k! \sum_{r_1=1}^{t} \frac{1}{r_1} \sum_{r_2=r_1+1}^{t} \frac{1}{r_2 - r_1} \ldots \sum_{r_k=r_{k-1}+1}^{t} \frac{1}{r_k - r_{k-1}} \\ &= k! \sum_{r_1=1}^{t} \frac{1}{r_1} \sum_{r_2=r_1}^{t} \frac{1}{r_2 - r_1 + 1} \ldots \sum_{r_{k-1}=r_{k-2}}^{t} \frac{1}{r_{k-2} - r_{k-1} + 1} \cdot O(\log t).\end{aligned}$$

We repeat this simplification for each level of summation replacing $\sum_{r_i=r_{i-1}}^{t} \frac{1}{r_i - r_{i-1} + 1}$ with $O(\log t)$. Iterating through the k levels gives $\mathbb{E}\left[c_j^k\right] = O(k! \log^k t)$ giving the lemma. $\square$