The Range of Random Walk Bridges

By Marios Stamoulis
Supervised by Anđela Šarković

Acknowledgements

Many thanks to **Anđela Šarković** and **Thomas Sauerwald** for their help in this project!

Another thank-you to the **Gatsby Foundation** and all of the other donors who funded this project and many more in the King's Summer Research Programme!

(This project was a part of the King’s Summer Research programme)

Presentation Outline

- Preliminaries
- Motivations
- Counting subgraphs
- Intro to Expanders
- *The Result*
- Proof (if we have time)

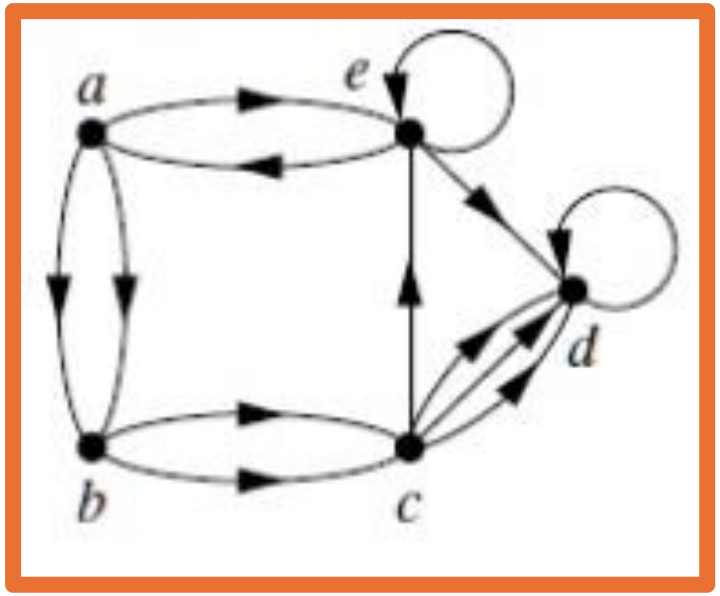
Antony Gormley's Quantum Cloud sculpture in London was designed by a computer using a random walk algorithm

A **Random Walk Bridge** is a **Closed Random Walk:**

A walk which is **conditioned to return to its start vertex** at time **t**

(next vertex chosen uniformly)

In this project, we use a **simple random walk** which traverses a finite **bounded-degree graph** (where every vertex has **at most Δ neighbours**)



A directed graph with loops and multi-edges

We assume that the graph is **undirected** and has **no loops or multi-edges**

Why these prerequisites?

Simple random walk → Avoids extreme cases

Finite, bounded degree → Useful bounds

(on transition probabilities and
the stationary distribution)

Why this project?

Random walks have uses in countless algorithms[1]

- Determining Satisfiability
- Estimating information spreading
- Approximating the volume of convex bodies

1. Thomas Sauerwald, He Sun, and Danny Vagnozzi. 'The Support of Open Versus Closed Random Walks'. doi: 10.4230/LIPIcs.ICALP.2023.103

- Determining Connectedness
- Estimating network sizes and densities
- The voter model
- Graph exploration
- Analysis of Randomness Amplification
- Estimating load balancing
- Electrical networks
- Other things in geometry, group theory, etc.

Motivations continued…

Open random walks have many known results

There is currently far less research on **closed** random walks

From a paper[1] on **closed** random walks…

"…it is tempting to conjecture that on any bounded-degree **expander** graph, the lower bound on the [**range**] can be improved, possibly even to **$\Omega(t)$**, which would … match the bound for open random walks"

1. Thomas Sauerwald, He Sun, and Danny Vagnozzi. 'The Support of Open Versus Closed Random Walks'. doi: 10.4230/LIPIcs.ICALP.2023.103

The **Range** of a Random Walk is the number of unique vertices it visits

The **range** of a walk is **at most linear** in the length of the walk

This means the range's long-term behaviour is at least linear in t

Therefore, a lower bound of **Ω(t)** could not be improved upon

We'll cover what 'expanders' are later…

Counting Subgraphs

A result we'll use later

Later we will need to know an upper bound for the **number of connected sets of size s containing some fixed start vertex**

In this section, we will find such a bound

This method was found in the following paper;
Theo McKenzie, Peter M. R. Rasmussen, and Nikhil Srivastava. 'Support of Closed Walks and Second Eigenvalue Multiplicity of the Normalized Adjacency Matrix'. url: https://arxiv.org/abs/2007.12819

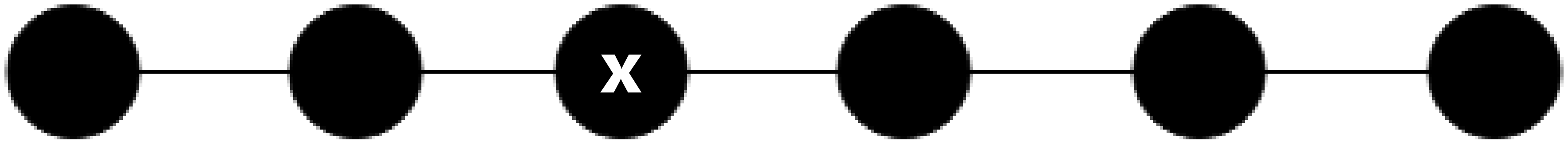
We may be tempted to find a surjection like this;

Random walks of length *s* starting from *x* $\twoheadrightarrow$ **Connected sets of size *s* around *x***

(we send a walk of length s to the set of vertices it visited)

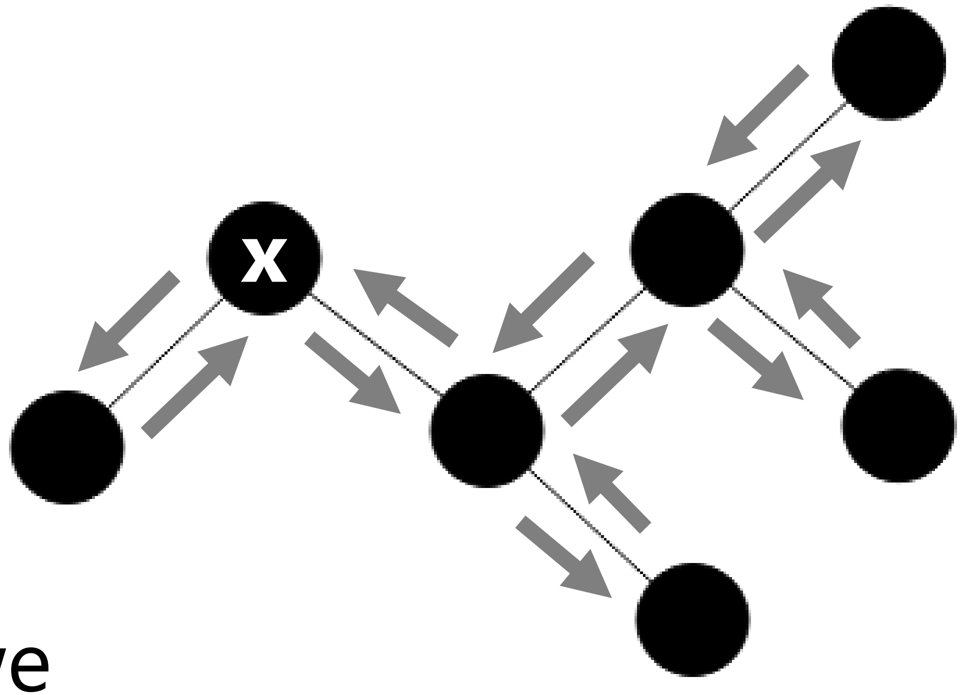
There are at most $\mathbf{\Delta^s}$ random walks of length *s* starting from *x*

However, not all of these subgraphs can be covered by a random walk of length *s*



See that a walk of length 6 starting at *x* **cannot** cover this subgraph of size 6

In **2s** steps, we can travel along **each edge in both directions**, ending at *x* again

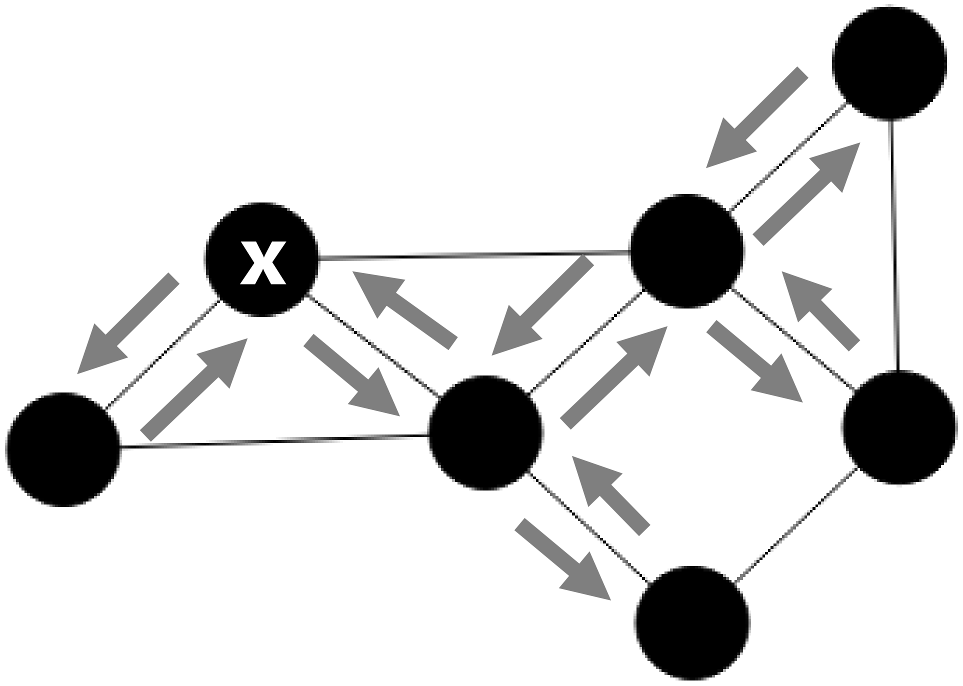


For any subgraph of size s, we will hit all vertices this way

(in this example, it looks like we're making an outline of the set)

There are at most $\boldsymbol{\Delta^{2s}}$ random walks of length $2s$ starting from *x*

(This method may *seem* to only work for 'tree-like' subgraphs



but we can either consider the subgraph's **spanning tree** or remove edges until the subgraph is **minimally connected** to make it look 'tree-like')

(if you can already see why this method works for any subgraph of size **s** then this slide can be skipped!)

Therefore, our surjection instead reads as

Random walks of length <u>2*s*</u> starting from *x*	$\twoheadrightarrow$	**Connected sets of size *s* around *x***

(we send a walk of length <u>2s</u> to the set of vertices it visited)

Hence there are at most $\mathbf{\Delta^{2s}}$ subgraphs of size *s* around *x*

Expander Graphs

And the Relaxation Time

Our result relies on the following Lemma[1]:

For a non-empty subset A and non-negative t;

$$\mathbb{P}_\pi(T_A > t) \leq \pi\left(A^c\right)\exp\left(-\frac{t \cdot \pi\left(A\right)}{t_{\mathrm{rel}}}\right)$$

where $T_A = \inf\{t \geq 0 : X_t \in A\}$ *is the first hitting time of the set* A *and the relaxation time is* $t_{\mathrm{rel}} = \frac{1}{1-\lambda_2}$ *where* λ_2 *is the second largest eigenvalue of the transition matrix* P*. Note that, for a family of expander graphs,* t_{rel} *is uniformly bounded.*

1. Riddhipratim Basu, Jonathan Hermon, and Yuval Peres. 'Characterization of cutoff for reversible Markov chains'. url: http://dx.doi.org/10.1214/16-AOP1090.

$$\mathbb{P}_\pi(T_A > t) \leq \pi(A^c) \exp\left(-\frac{t \cdot \pi(A)}{t_{\mathrm{rel}}}\right)$$

We use this to **bound the probability of staying in a set**

(Since **leaving A** is the same as **hitting A^c**)

Now, let's discuss **expander graphs**;

Expander Graphs satisfy the following;

In any small subset A of the graph, the number of **edges leaving the subset** is large - at least **c|A|** for a fixed constant c

In an expander graph, **it is hard to stay within a small subset** because the graph is so well-connected

Studying expanders is useful because **randomly-generated** graphs are likely to be expanders

If we uniformly choose a d-regular graph with n vertices and choose some small enough $\alpha>0$ then the probability the graph will be an "α-expander" tends to 1 as n tends to infinity[1]

Additionally, they have bounds on their **relaxation time, t_{rel}** – this will be useful later!

1. Nathanaël Berestycki. 'Mixing Times of Markov Chains: Techniques and Examples'. url: https://personal.math.ubc.ca/~jhermon/Mixing/mixing3.pdf.

The result we found in this project is **especially strong for expander graphs**

It proves the **tight bound** mentioned by **Sauerwald et. al** for large enough walk-lengths

A Bound for the Range of Random Walk Bridges

In relation to the Relaxation Time

(The Result)

Below is the result in full – don't worry about reading the fine print!

Theorem 1.2. *For all $\Delta > 0$ there is a large enough constant $C > 0$ such that for any graph $G = (V, E)$ with degrees bounded by Δ and $|V| = n$, for all even times t satisfying $\frac{n}{2} \geq t \geq C\, t_{\mathrm{rel}} \log n$, where t_{rel} is the relaxation time of the simple random walk on G, we have that for all sufficiently small c there is a constant c' (depending on c) such that*

$$\mathbb{P}_x\left(R_t > \frac{ct}{t_{\mathrm{rel}}} \,\middle|\, X_t = x \right) \geq 1 - e^{-c' \frac{t}{t_{\mathrm{rel}}}}$$

where X_t is the simple random walk on G and R_t is its range.

The main strategy of the proof is to **bound the probability that a walk has at most range s at time *t*** by the probability that a walk of length *t* **stays inside some connected set of size *s*** containing the start vertex

Later, we'll replace **s** with something more useful

$$\mathbb{P}_x(R_t < s) \leq \sum_{A} \mathbb{P}_x(T_{A^c} > t)$$

(we use the union bound here)

We found that the **number of connected subsets of size s** is **upper-bounded by Δ^{2s}**

$$\sum_{A} \mathbb{P}_x(T_{A^c} > t) \leq \Delta^{2s} \max_{A} \mathbb{P}_x(T_{A^c} > t)$$

We then recall the lemma from earlier

$$\mathbb{P}_\pi(T_A > t) \leq \pi(A^c)\exp\left(-\frac{t \cdot \pi(A)}{t_{\text{rel}}}\right)$$

With some tampering…

$$\Delta^{2s} \max_A \mathbb{P}_x(T_{A^c} > t) \leq \Delta^{2s-1} s \cdot \exp\left(-\frac{t}{2\Delta \cdot t_{\text{rel}}}\right)$$

(there's a **lot** that I'm glossing over in this step)

We can then lazily re-introduce our conditioning and find that the **right-hand side tends to 0 in certain conditions**

$$\mathbb{P}_x(R_t < s \mid X_t = x) \leq \frac{\mathbb{P}_x(R_t < s)}{\mathbb{P}_x(X_t = x)} \leq \Delta^{2s} sn \cdot \exp\left(-\frac{t}{2\Delta \cdot t_{\mathrm{rel}}}\right)$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A,B)}{\mathbb{P}(B)} \leq \frac{\mathbb{P}(A) \cdot \mathbb{P}(B|A)}{\mathbb{P}(B)} \leq \frac{\mathbb{P}(A)}{\mathbb{P}(B)}$$

We set $s = c\,\frac{t}{t_{\mathrm{rel}}}$ and find there exists some constant

C > 0 such that the right-hand side tends to 0 for **$t \geq C\, t_{rel} \log n$**

$$\Delta^{2s} sn \cdot \exp\left(-\frac{t}{2\Delta \cdot t_{\mathrm{rel}}}\right)$$

Therefore, the **probability that a closed random walk has range below ct/t_{rel} tends to 0** for **$t \geq C\, t_{rel} \log n$**

$$\mathbb{P}_x\left(R_t < c \cdot \frac{t}{t_{\mathrm{rel}}} \;\middle|\; X_t = x \right) \to 0$$

This concludes the proof.

□

Remember expanders?

In a family of expanders, the **relaxation time, t_{rel}** is bounded

In the case of **expander graphs**, our result is strengthened. **For t ≥ C log n**,

$$\mathbb{P}_x(R_t < c \cdot t \mid X_t = x) \to 0$$

This is a **tight bound** – it cannot be improved upon!

This concludes the presentation!

Thank you for listening!

If you want any more information, feel free to send an email to me at ms2911@cam.ac.uk.