Coverage probabilities of edges in a regular graph by a random walk

Zheng Choo

Keble College

University of Oxford

A thesis submitted in fulfillment of the requirements for the degree of Master of Science by Research

August 2014

Abstract

We investigate the probability that a random walk covers a given edge in a connected regular network. For a complete graph or a circle graph, we obtain an explicit expression of the coverage probability by using the walk generating function. Alternatively, for any fixed connected regular network, we use the largest eigenvalue of the adjacency matrix of the network with the given edge removed from the original network to give a lower bound for the coverage probability. Finally, we give the results by Cooper and Frieze [12] on the coverage probabilities of nice edges for random regular networks.

The motivation of this project is to obtain properties of a network from a new perspective, which is through primes of a network using random walks. Primes are simple paths and simple cycles in a network, which can be seen as building blocks of a network. A network is uniquely determined by its set of primes.

A network can be explored through a random walk on it, and the random walk can be used to detect primes of the network. A necessary condition for retrieving all the primes in a network is that a random walk visits all edges in the network. The approach of investigating the coverage probability of a given edge serves as a first step to investigate the coverage probabilities of primes.

Acknowledgements

I am very grateful that my supervisors, Professor Gesine Reinert and Professor Dieter Jaksch, taught me a lot of precious skills in doing research. Their enthusiasm for research was inspiring and led me to endless wonder. I would really like to thank them for revising this thesis with patience.

I would like to thank Dr Pierre-Louis Giscard for his useful ideas and valuable suggestions from our discussions during the period of this project. I would also like to acknowledge Professor Colin Cooper for his constructive and insightful suggestions.

List of Notations

$w^{\mu\nu}_{\mathcal{G};\ell}$	a walk of length ℓ on $\mathcal{G}$ from μ to ν
$w^{\mu}_{\mathcal{G};\ell}$	a walk of length ℓ on $\mathcal{G}$ starting from μ
$\mathcal{E}(w^{\mu\nu_\ell}_{\mathcal{G};\ell})$	the set of edges visited by $w^{\mu\nu}_{\mathcal{G};\ell}$
$\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell}$	the set of all walks of length ℓ on $\mathcal{G}$ from μ to ν
$\mathfrak{W}^{\mu}_{\mathcal{G};\ell}$	the set of all walks of length ℓ on $\mathcal{G}$ starting from μ
$\mathfrak{W}_{\mathcal{G};\ell}$	the set of all walks of length ℓ on $\mathcal{G}$
$\mathcal{P}_{\mathcal{G}}$	the set of all simple paths on $\mathcal{G}$
$\mathcal{C}_{\mathcal{G}}$	the set of all simple cycles on $\mathcal{G}$
$\mathcal{G} \setminus \{\mu_1, \mu_2, \ldots, \mu_i\}$	the subgraph of $\mathcal{G}$ with vertices $\{\mu_1, \mu_2, \ldots, \mu_i\} \subset \mathcal{V}$ and all edges incident to them removed from $\mathcal{G}$
$\mathcal{G} \setminus \{e\}$	the subgraph of $\mathcal{G}$ with the edge $e \in \mathcal{E}$ removed from $\mathcal{G}$
$\mathcal{G}_r$	a fixed connected r-regular graph
$\mathcal{C}_n$	a circle graph on n vertices
$\mathcal{K}_n$	the complete graph on n vertices
$\mathcal{F}^{\mu}_{\mathcal{G};\ell}$	the power set of $\mathfrak{W}^{\mu}_{\mathcal{G};\ell}$
$m(\cdot)$	the counting measure on $\mathcal{F}^{\mu}_{\mathcal{G};\ell}$
$\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\cdot)$	a probability measure on $\mathcal{F}^{\mu}_{\mathcal{G};\ell}$

$\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}$	the event that a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ visits a given edge $e \in \mathcal{E}(\mathcal{G})$
$\mathcal{W}^{\mu}_{\mathcal{G};\ell}$	a random walk of length ℓ on $\mathcal{G}$, starting from μ
$g^{\mu\nu}_{\mathcal{G}}(z)$	the walk generating function associated with $\{m(\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell})\}_{\ell\geq 0}$
$g^{\mu}_{\mathcal{G}}(z)$	the walk generating function associated with $\{m(\mathfrak{W}^{\mu}_{\mathcal{G};\ell})\}_{\ell\geq 0}$
$A_{\mathcal{G}}$	the adjacency matrix of $\mathcal{G}$
$\mathcal{G}(n,k)$	k-nearest-neighbour graph
$\lambda_{max}(A)$	the largest eigenvalue of A
$\mathfrak{G}^{(r)}_{n}$	the set of simple r-regular graphs on n vertices
$\mathcal{G}^{rnd}_{r}$	a random r-regular graph
$\odot$	nesting product
$S(w)$	the set of prime factors of a walk w
$\chi^{\alpha\alpha'}_{\mathcal{G}}$	a prime on $\mathcal{G}$ from α to α'

Contents

1 Introduction

Networks arise in a large variety of realms such as social science, physics, biology, finance and computer science, and many real world datasets can be visualised as networks to provide intuition for an investigation of the datasets (see [33] for examples). A network is a graph $\mathcal{G} = (\mathcal{V}(\mathcal{G}), \mathcal{E}(\mathcal{G}))$ comprising a set of vertices $\mathcal{V}(\mathcal{G})$ and a set of edges $\mathcal{E}(\mathcal{G})$. We shall write $\mathcal{V} = \mathcal{V}(\mathcal{G})$ and $\mathcal{E} = \mathcal{E}(\mathcal{G})$ when it is clear from the context what $\mathcal{G}$ is. We follow the common parlance of using the words "network" and "graph" interchangeably.

Vertices in a network may correspond to different objects and edges represent connections between them. For instance, in a social network, vertices can represent individuals and an edge is created between two individuals if they are friends. Edges in a network can be directed or undirected, yielding a directed graph (digraph) or an undirected graph respectively. We shall only consider undirected graphs unless specified otherwise. The existence of an edge between two vertices can be random, i.e. we put probabilities on edges which are given by some law. In this case, $\mathcal{G}$ is called a random graph. If the existence of an edge between two vertices is deterministic, we call the graph $\mathcal{G}$ a deterministic graph or a fixed graph.

On a fixed graph $\mathcal{G}$, the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ can be used to calculate some properties of $\mathcal{G}$. For $\ell \geq 2$, the (μ, ν) entry of $A_{\mathcal{G}}^{\ell}$, where $\mu, \nu \in \mathcal{V}$, is the number of walks of length ℓ from μ to ν. When $\ell = 3$, the (μ, μ) entry of $A_{\mathcal{G}}^3$ gives the number of walks of length 3 from μ to itself. The number of triangles in $\mathcal{G}$ is given by $\frac{1}{6}\text{trace}(A_{\mathcal{G}}^3)$ [8]. The minimum value $\ell \geq 2$ such that the (μ, ν) entry of $A_{\mathcal{G}}^{\ell}$, where $\mu \neq \nu$, is nonzero gives the length of the shortest path (not necessarily unique) from μ to ν. Other properties of a fixed graph $\mathcal{G}$ such as subgraph centrality [15] and communicability [16] are defined in terms of the matrix exponential $e^{A_{\mathcal{G}}}$. The subgraph centrality of a vertex μ is the (μ, μ) entry of $e^{A_\mathcal{G}}$, while the communicability between μ and ν, where $\mu \neq \nu$, is the (μ, ν) entry of $e^{A_\mathcal{G}}$. Notice that $e^{A_\mathcal{G}}$ is a power series of $A_\mathcal{G}$. For a real-world network consisting of over 10 million vertices (for example the World Wide Web), calculating matrix powers of $A_\mathcal{G}$ is computationally expensive. Hence approximations can be useful.

One such approximation is through random walks on graphs. A network can be explored through a random walk or random walks on it. For instance, Costa and Travieso [14] used random walks to estimate the average vertex degrees and the average clustering coefficients of Bernoulli random graphs and Barabási-Albert network models. Cooper *et al.* [13] considered how to use random walks to estimate the number of vertices, edges and triangles in finite connected and undirected graphs. They determined the expected value of each property and give the accuracy of their estimate. Asztalos and Toroczkai [3] gave an exact formula for the average number of unique, visited edges as a function of time for arbitrary (deterministic) graphs. The authors also introduced the edge explorer model (EEM) which performs edge discovery well on networks. Cooper and Frieze [12] gave the component structure of the set of unexplored vertices in terms of the exploring time of random walks on Bernoulli random graphs, random regular graphs and random digraphs.

In this thesis, we investigate the probability that a random walk visits a given edge in a connected regular graph. The motivation of this project is to obtain network properties from a new perspective, that is, through *primes* of a network by using *random walks*. Primes are *simple paths* and *simple cycles* in a graph [19], which can be seen as building blocks of a graph. This concept is akin to what prime numbers are to integers [19]. A simple path is a sequence of vertices which are all distinct. A simple cycle is a sequence of vertices whose endpoints are identical and intermediate vertices (from the second to the penultimate) are all distinct and different from the endpoints. Edges in a network are special cases of primes. Thus any network without isolated vertices is uniquely determined by its set of primes.

In terms of applications, we give a sense of what can be done after obtaining the set of *all* primes in a graph. In [20], Giscard *et al.* introduced the method of path-sums to evaluate matrix functions in a *finite* number of steps. They showed that for any square matrix A, any matrix function $f(A)$ can be evaluated as a finite sum over primes. The idea is to partition A into submatrices and associate them with a weighted directed graph $\mathcal{G}$. Then evaluating any submatrix of $f(A)$ is equivalent to summing the weights of all primes between a particular pair of vertices in $\mathcal{G}$. For any finite square matrix, the corresponding directed graph $\mathcal{G}$ is finite. Consequently, the number of primes between any pair of vertices in $\mathcal{G}$ is finite. The path-sums method turns, for example, the problem of evaluating e^A, which is a power series, into evaluating a finite sum over primes. We recall that $e^{A_\mathcal{G}}$ is related to subgraph centrality and communicability of $\mathcal{G}$. We can then use the set of all primes in $\mathcal{G}$ and the path-sum method to evaluate $e^{A_\mathcal{G}}$ and obtain these two network properties.

Assume that we have the set of all primes $\chi(\mathcal{G})$ in a network $\mathcal{G}$. We can use this set of primes to reveal some properties of a network, such as the following.

(a) Shortest paths: From the definition of primes given above, shortest paths between μ and ν, where $\mu, \nu \in \mathcal{V}$, are primes. One can obtain the representation and the length of the shortest paths from $\chi(\mathcal{G})$.

(b) Diameter: Similarly, one can obtain the diameter of $\mathcal{G}$ by comparing the shortest primes between any pair of vertices in $\chi(\mathcal{G})$.

(c) Subgraph counts: Another possibility is to use $\chi(\mathcal{G})$ to reveal fixed subgraphs in $\mathcal{G}$. Subgraphs of fixed patterns occurring in a network are either primes by themselves or combinations of several short primes.

Thus, it suffices to obtain *all* primes in a network $\mathcal{G}$ in order to study $\mathcal{G}$. In a similar spirit to using random walks to explore networks, we employ the idea of using random walks to recover primes in networks. A *necessary* condition for getting all primes in $\mathcal{G}$ is that a random walk visits all edges, which are special cases of primes, in $\mathcal{G}$. Obtaining the probability of a random walk visiting a given edge serves as a first step of investigating the probabilities of visiting a given prime. To obtain the *representations* of primes from a walk, we use the prime factorisation method developed by Giscard *et al.* in [19].

This thesis is organised as follows. In Section 2, we introduce some notations needed throughout the thesis and properties of graphs which we consider. We introduce the random walk model that we use throughout the thesis in Section 3. In Section 4, we discuss the difference between the problem of missing an edge in a random walk and the problem of missing a vertex in a random walk, and show a connection between these two problems by using the notion of an edge-vertex dual graph. In Section 5, we obtain an explicit expression of the coverage probability of a given edge for a complete graph or a circle graph by using the walk generating function. Alternatively, for any fixed connected regular graph, we provide a lower bound for the coverage probability of a given edge by using the largest eigenvalue of the adjacency matrix of the graph with the given edge removed from the original graph in Section 6. In Section 7, we give recent results by Cooper and Frieze [12] on the coverage probabilities of nice edges for a random regular graph. On the recoverability of primes from a walk, we introduce the prime factorisation method in Section 8. Finally, we conclude this thesis with some exploratory simulations and open problems.

2 Notations and properties of graphs

A walk $w^{\mu\nu}_{\mathcal{G};\ell}$ of length ℓ on $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ from μ to ν, where $\mu, \nu \in \mathcal{V}$, is a left-to-right sequence of vertices $(\mu\mu_1 \dots \mu_{\ell-1}\nu)$ in which $\{(\mu_i, \mu_{i+1}) : i = 0, 1, \dots, \ell - 1\} \subset \mathcal{E}$ with $\mu_0 \equiv \mu$ and $\mu_\ell \equiv \nu$. We write $w^{\mu}_{\mathcal{G};\ell}$ to denote a walk of length ℓ on $\mathcal{G}$ starting from μ. We denote the set of edges visited by a walk $w^{\mu\nu}_{\mathcal{G};\ell}$ by $\mathcal{E}(w^{\mu\nu}_{\mathcal{G};\ell})$. The set of all walks of length ℓ on $\mathcal{G}$ from μ to ν is denoted by $\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell}$. The set of all walks of length ℓ on $\mathcal{G}$ starting from μ is denoted by $\mathfrak{W}^{\mu}_{\mathcal{G};\ell} = \cup_{\nu\in\mathcal{V}} \mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell}$ and the set of all walks of length ℓ on $\mathcal{G}$ is denoted by $\mathfrak{W}_{\mathcal{G};\ell} = \cup_{\mu\in\mathcal{V}} \mathfrak{W}^{\mu}_{\mathcal{G};\ell}$.

A walk $w^{\mu\nu}_{\mathcal{G};\ell}$ is called a path when $\mu \neq \nu$; otherwise, it is called a cycle. We recall that a simple path is one in which all vertices are distinct. A simple cycle is a cycle whose intermediate vertices (from the second to the penultimate) are all distinct and are different from the endpoints. We denote the set of all simple paths on $\mathcal{G}$ by $\mathcal{P}_\mathcal{G}$ and the set of all simple cycles on $\mathcal{G}$ by $\mathcal{C}_\mathcal{G}$.

We denote by $\mathcal{G} \setminus \{\mu_1, \mu_2, \dots, \mu_i\}$ the subgraph of $\mathcal{G}$ with vertices $\{\mu_1, \mu_2, \dots, \mu_i\} \subset \mathcal{V}$ and all edges incident to them removed from $\mathcal{G}$. We shall also write $\mathcal{G} \setminus \{e\}$ to denote the subgraph of $\mathcal{G}$ with the edge $e \in \mathcal{E}$ removed from $\mathcal{G}$.

We consider fixed graphs and random graphs. For simplicity, we consider fixed graphs $\mathcal{G}$ with the following properties:

(P1) $\mathcal{G}$ is finite, i.e. the number of vertices $|\mathcal{V}| = n$ and the number of edges $|\mathcal{E}| = m$ are finite.

(P2) $\mathcal{G}$ is undirected, i.e. for any pair of vertices $\mu, \nu \in \mathcal{V}$, $(\mu, \nu) \in \mathcal{E}$ implies $(\nu, \mu) \in \mathcal{E}$.

(P3) Multiple edges and self-loops are not allowed in the graph, i.e. $\mathcal{G}$ is a simple graph.

(P4) $\mathcal{G}$ in unweighted, i.e. all edges have weight 1.

(P5) $\mathcal{G}$ is connected, i.e. for any pair of vertices $\mu, \nu \in \mathcal{V}$ and for some $\ell \geq 1$, there exists a walk $w^{\mu\nu}_{\mathcal{G};\ell}$ from μ to ν.

Note that Property 5 ensures the existence of at least one simple path between every pair of vertices in $\mathcal{G}$. When we consider random graphs, we assume that (P5) holds *with high probability* (see Section 7 for details).

A graph $\mathcal{G}$ is called an r-regular graph if each vertex in $\mathcal{G}$ has degree r. In this thesis, we consider r-regular graphs on n vertices, where nr is even. As the sum of the degrees of vertices in $\mathcal{G}$ is twice the number of edges, for n odd, an r-regular graph exists only when r is even. From now on, we denote a fixed connected r-regular graph by $\mathcal{G}_r$. When $r = 2$, $\mathcal{G}_r$ is a circle graph on n vertices, which is denoted by $\mathcal{C}_n$. When $r = n - 1$, $\mathcal{G}_r$ is the complete graph on n vertices, which is denoted by $\mathcal{K}_n$.

3 Random walks on graphs

Given a starting point μ, a random walk on a graph $\mathcal{G}$ proceeds by selecting one of the neighbours of μ uniformly at random (u.a.r.) and moving to the vertex; then the walk selects one of the neighbours of this vertex u.a.r. and moves to it, and so on. From now on, we say "a random walk" to mean a random walk generated by this procedure.

Let $\mathcal{G} = \mathcal{G}_r$ and $\mathcal{F}^{\mu}_{\mathcal{G};\ell}$ be the power set of $\mathfrak{W}^{\mu}_{\mathcal{G};\ell}$. Let $m(\cdot)$ be the counting measure on $\mathcal{F}^{\mu}_{\mathcal{G};\ell}$, i.e. $m(S)$ is the number of elements in $S \in \mathcal{F}^{\mu}_{\mathcal{G};\ell}$. We define the probability space of walks of length ℓ, starting from μ, as $\left(\mathfrak{W}^{\mu}_{\mathcal{G};\ell}, \mathcal{F}^{\mu}_{\mathcal{G};\ell}, \mathbb{P}^{\mu}_{\mathcal{G};\ell}\right)$ with a probability measure $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\cdot)$ defined as $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(S) = \frac{m(S)}{r^\ell}$ $(*)$, where $S \in \mathcal{F}^{\mu}_{\mathcal{G};\ell}$. Lemma 1 below shows $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\cdot)$ is well-defined.

Lemma 1. *For a fixed r-regular graph $\mathcal{G}_r$, where $r \geq 2$, we have $m(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}) = r^\ell$. In particular, $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\cdot)$ gives a probability measure on $\mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$.*

Proof. Starting from μ, there are r neighbours from which a walk of length ℓ can select at each step, thus we have $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}\right) = r^{\ell}$. Now we show that $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\cdot)$ gives a probability measure on $\mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$, where $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(S) = \frac{m(S)}{r^{\ell}}$ for $S \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$.

First of all, for $S \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$, we have $0 \leq m(S) \leq r^{\ell}$, thus $0 \leq \mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(S) \leq 1$. Then we have $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}\right) = \frac{m\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}\right)}{r^{\ell}} = 1$. Finally, for a countable collection of pairwise disjoint events $\{S_i : S_i \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}\}$,

$$\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}\left(\cup_{S_i \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}} S_i\right) = \frac{m\left(\cup_{S_i \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}} S_i\right)}{r^{\ell}} = \frac{\sum_{S_i \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}} m(S_i)}{r^{\ell}} = \sum_{S_i \in \mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}} \mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(S_i).$$

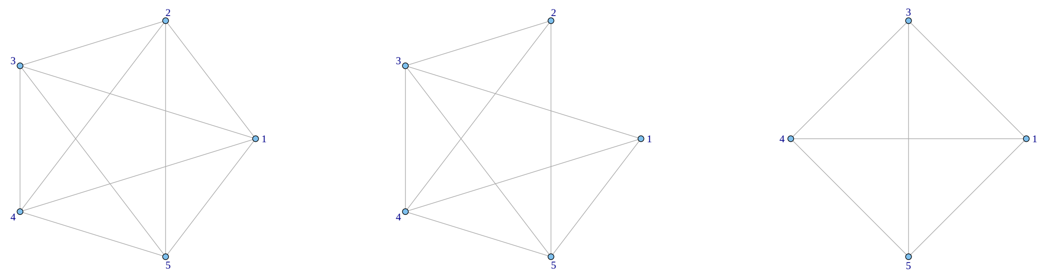
Here we use the countable additivity of the counting measure $m(\cdot)$ on $\mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$ in the second equality. Thus, $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\cdot)$ gives a probability measure on $\mathcal{F}^{\mu}_{\mathcal{G}_r;\ell}$. $\square$

Lemma 1 says that, starting from μ, each walk of length ℓ on a fixed regular graph is equally likely to be a realisation of a random walk of length ℓ, starting from μ. On a fixed graph which is not regular, a random walk does not necessarily result in all realisations of the walk being equally likely.

From now on, we denote a random walk of length ℓ on $\mathcal{G}$, starting from μ, by $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$. Using Lemma 1, we can relate the problem of finding the probability that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ misses a given edge $e \in \mathcal{E}(\mathcal{G})$ to a combinatorial problem: given a starting point μ, how many walks of length ℓ miss a given edge $e \in \mathcal{E}(\mathcal{G})$? We shall use this relation in Section 5.

4 Missing an edge vs. missing a vertex

We are interested in the probability that a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on a graph $\mathcal{G}$ misses *a given edge* $e \in \mathcal{E}$. The question we are looking into is different from obtaining the probability that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ misses a given vertex $\nu \in \mathcal{V}$, where $\nu \neq \mu$. To illustrate the difference, we give an example for the case when $\mathcal{G}$ is the complete graph $\mathcal{K}_5$ on 5 vertices (see Fig. 1a).



(a) (b) (c)

Figure 1: (a) The complete graph $\mathcal{K}_5$ on 5 vertices; (b) The graph $\mathcal{K}_5 \setminus \{(1,2)\}$ with edge $(1,2)$ removed; (c) The graph $\mathcal{K}_5 \setminus \{2\}$ with vertex 2 and all edges incident to it removed.

Example 1. Consider a random walk $\mathcal{W}^1_{\mathcal{K}_5;2}$ of length 2 on $\mathcal{K}_5$, starting from vertex 1.
(a) Missing an edge: The total number of walks in $\mathfrak{W}^1_{\mathcal{K}_5;2}$ that miss edge $(1,2)$ is 12 (by counting the number of walks of length 2 on the graph shown in Fig. 1b, starting from vertex 1) and $m\left(\mathfrak{W}^1_{\mathcal{K}_5;2}\right) = 16$. By Lemma 1, we obtain the probability that $\mathcal{W}^1_{\mathcal{K}_5;2}$ misses $(1,2)$, which is $\frac{12}{16} = \frac{3}{4}$.
(b) Missing a vertex: Similarly, since the total number of walks in $\mathfrak{W}^1_{\mathcal{K}_5;2}$ that miss vertex 2 is 9 (by counting the number of walks of length 2 on the graph shown in Fig. 1c, starting from vertex 1), the probability that $\mathcal{W}^1_{\mathcal{K}_5;2}$ misses vertex 2 is $\frac{9}{16}$, which is smaller than the probability that that $\mathcal{W}^1_{\mathcal{K}_5;2}$ misses $(1,2)$.

In general, for the complete graph $\mathcal{K}_n$ on n vertices, the probability that $\mathcal{W}^\mu_{\mathcal{K}_n;\ell}$ misses a vertex $\nu \neq \mu$ in $\mathcal{K}_n$ is $\left(1 - \frac{1}{n-1}\right)^\ell$ since at each step, the probability that the walk does not visit ν is $\left(1 - \frac{1}{n-1}\right)$ (see Example 2 in [32]). From Theorem 3 in Section 5, we shall see that there is a difference between the probability that $\mathcal{W}^\mu_{\mathcal{K}_n;\ell}$ misses a given vertex $\nu \neq \mu$ in $\mathcal{K}_n$ and the probability that $\mathcal{W}^{\mu}_{\mathcal{K}_n;\ell}$ misses a given edge e in $\mathcal{K}_n$.

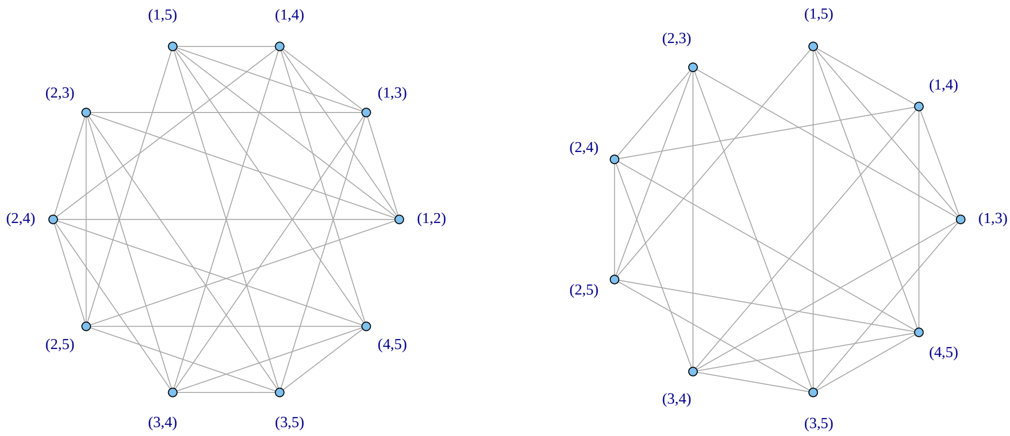
Let μ be one of the endpoints of a given edge e in $\mathcal{K}_n$. Let d_ℓ be the number of walks of length ℓ starting from μ which end at one of the endpoints of e, and let s_ℓ be the number of walks of length ℓ starting from μ which do not end at an endpoint of e. One might think that the following recurrence relations would solve the problem of obtaining the probability that a random walk of length ℓ starting from μ visits e,

$$s_\ell = (n-3)s_{\ell-1} + (n-2)d_{\ell-1}, \tag{1}$$

$$d_\ell = d_{\ell-1} + 2s_{\ell-1}, \tag{2}$$

with $s_0 = 0$ and $d_0 = 1$. The first recurrence relation (1) comes from the fact that on $\mathcal{K}_n$, a walk of length $\ell - 1$ that does not end at an endpoint of e has $n-3$ vertices to move on without ending at an endpoint of e, and a walk of length $\ell - 1$ that ends at one of the endpoints of e can move away from the endpoint of e by selecting one of the $n-2$ vertices. Similarly, the second recurrence relation (2) comes from the fact by taking one more step along e, a walk of length $\ell - 1$ that ends at one of the endpoints of e can visit another endpoint of e, and a walk of length $\ell - 1$ that does not end at an endpoint of e can visit μ or another endpoint of e by taking one more step. However, the problem of the recurrence relations is that visiting one of the endpoints of e does not mean that e is visited. In fact, both s_ℓ and d_ℓ take the walks that visit e *and* those which do not visit e into account.

The problem of a random walk missing an edge can be related to the problem of a random walk missing a vertex through the notion of an *edge-vertex dual graph*. Given a graph $\mathcal{G}$, the edge-vertex dual graph $\mathcal{G}'$ of $\mathcal{G}$, which is also called the line graph, is constructed as follows: i) for each edge in $\mathcal{G}$ draw a vertex in $\mathcal{G}'$; and ii) place an edge between two vertices in $\mathcal{G}'$ if the edges in $\mathcal{G}$ corresponding to these two edges are adjacent in $\mathcal{G}$. The edge-vertex dual graph of $\mathcal{K}_5$ is shown in Fig. 2a. Note that the edge-vertex dual graph $\mathcal{K}_5'$ of $\mathcal{K}_5$ is not a complete graph.



(a) (b)

Figure 2: (a) The edge-vertex dual graph $\mathcal{K}_5'$ of $\mathcal{K}_5$; (b) The graph $\mathcal{K}_5' \setminus \{(1,2)\}$ with vertex $(1,2)$ and all edges incident to it removed.

Lemma 2. *For an r-regular graph $\mathcal{G} = (\mathcal{V}(\mathcal{G}), \mathcal{E}(\mathcal{G}))$ on n vertices, its edge-vertex dual graph $\mathcal{G}' = (\mathcal{V}(\mathcal{G}'), \mathcal{E}(\mathcal{G}'))$ is a $(2r-2)$-regular graph on $\frac{nr}{2}$ vertices.*

Proof. There are $(2r-2)$ edges which are adjacent to each edge $e \in \mathcal{E}(\mathcal{G})$. To construct the edge-vertex dual graph $\mathcal{G}'$ of $\mathcal{G}$, we represent each edge $e \in \mathcal{E}(\mathcal{G})$ as a vertex $v_e \in \mathcal{V}(\mathcal{G}')$ and an edge is created between two vertices in $\mathcal{G}'$ if the edges in $\mathcal{G}$ corresponding to these two vertices share an endpoint. Thus v_e has degree $(2r-2)$. Applying this to all edges in $\mathcal{G}$, we obtain $\mathcal{G}'$ which is a $(2r-2)$-regular graph. Note that $|\mathcal{E}(\mathcal{G})| = \frac{nr}{2}$ and $|\mathcal{V}(\mathcal{G}')| = |\mathcal{E}(\mathcal{G})|$. $\square$

Let $v_e \in \mathcal{V}(\mathcal{G}')$ be the vertex corresponding to a given edge $e \in \mathcal{E}(\mathcal{G})$. Unless $\mathcal{G}'$ is $\mathcal{K}_3$ or a complete bipartite graph $\mathcal{K}_{1,3}$ (a star graph with 3 edges), there is a one-to-one correspondence between $\mathcal{G}$ and $\mathcal{G}'$ (see Theorem 8.3 in [21]). Hence there is a link between the probability that a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$ misses e and the probability that a random walk $\mathcal{W}^{\mu'}_{\mathcal{G}';\ell'}$ on $\mathcal{G}'$ misses v_e. However, for such a link, it is not obvious how μ relates to μ' and how ℓ relates to ℓ'.

One way is to choose μ and μ' u.a.r. on $\mathcal{G}$ and $\mathcal{G}'$ respectively. Alternatively, one can also choose the starting point from a set of vertices $S \subset \mathcal{V}(\mathcal{G}')$ u.a.r. such that the edges in $\mathcal{G}$ corresponding to the elements of S share μ as an endpoint. We show by examples that, in each case, the probability that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$ misses e is generally *not* the same as the probability that $\mathcal{W}^{\mu'}_{\mathcal{G}';\ell'}$ on $\mathcal{G}'$ misses v_e, regardless of whether ℓ' equals ℓ or $\ell-1$ (a consideration that a walk of length ℓ on $\mathcal{G}$ corresponds to a walk of length $\ell-1$ on $\mathcal{G}'$ is reasonable since each edge in $\mathcal{G}$ is represented by a vertex in $\mathcal{G}'$).

Example 2. Consider a random walk $\mathcal{W}^{\mu}_{\mathcal{K}_5;2}$ on $\mathcal{K}_5$ in Fig. 1a and a random walk $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;\ell'}$ on the edge-vertex dual graph $\mathcal{K}'_5$ shown in Fig. 2a, where μ is chosen u.a.r from $\mathcal{V}(\mathcal{K}_5)$ and μ' is chosen u.a.r. from $\mathcal{V}(\mathcal{K}'_5)$.

(a) On $\mathcal{K}_5$: Since there are 66 walks in $\mathfrak{W}_{\mathcal{K}_5;2}$ that miss edge $(1,2)$ (by counting the number of walks of length 2 on the graph shown in Fig. 1b), the probability that $\mathcal{W}^{\mu}_{\mathcal{K}_5;2}$ misses $(1,2)$ is $\frac{1}{5} \times \left(\frac{1}{4}\right)^2 \times 66 = \frac{33}{40}$.

(b) On $\mathcal{K}'_5$: When $\ell' = 1$, there are 48 walks in $\mathfrak{W}_{\mathcal{K}'_5;1}$ that miss vertex $(1,2)$ (by counting the number of walks of length 1 on the graph shown in Fig. 2b). The probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;1}$ misses $(1,2)$ is $\frac{1}{10} \times \frac{1}{6} \times 48 = \frac{4}{5}$. When $\ell' = 2$, there are 258 walks in $\mathfrak{W}_{\mathcal{K}'_5;2}$ that miss vertex $(1,2)$ (by counting the number of walks of length 2 on the graph shown in Fig. 2b). The probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;2}$ misses $(1,2)$ is $\frac{1}{10} \times \frac{1}{36} \times 258 = \frac{43}{60}$.

Example 3. Now we choose the starting point of $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;\ell'}$ from a set of vertices $S \subset \mathcal{V}(\mathcal{G}')$ u.a.r. such that the edges in $\mathcal{G}$ corresponding to the elements of S share μ as an endpoint. Without loss of generality, set $\mu = 1$. From Example 1a, we know that the probability that $\mathcal{W}^1_{\mathcal{K}_5;2}$ misses $(1,2)$ is $\frac{3}{4}$.

(a) μ' **is chosen u.a.r. from** $S = \{(1,2),(1,3),(1,4),(1,5)\}$: We count the corresponding number of walks similarly as in the previous examples. When $\ell' = 1$, there are 15 walks in $\cup_{\mu' \in S} \mathfrak{W}^{\mu'}_{\mathcal{K}'_5;1}$ that miss vertex $(1,2)$. The probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;1}$ misses $(1,2)$ is $\frac{1}{4} \times \frac{1}{6} \times 15 = \frac{5}{8}$. When $\ell' = 2$, there are 81 walks in $\cup_{\mu' \in S} \mathfrak{W}^{\mu'}_{\mathcal{K}'_5;2}$ that miss vertex $(1,2)$. The probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;2}$ misses $(1,2)$ is $\frac{1}{4} \times \frac{1}{36} \times 81 = \frac{9}{16}$.

(b) Not revisiting $(1,2)$: If μ' is chosen to be $(1,2)$, we can consider the question of not revisiting $(1,2)$. Then the probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;1}$ misses $(1,2)$ is still $\frac{1}{4} \times \frac{1}{6} \times 15 = \frac{5}{8}$ but the probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;2}$ misses $(1,2)$ becomes $\frac{1}{4} \times \frac{1}{36} \times 111 = \frac{37}{48}$.

(c) μ' **is chosen u.a.r. from** $S \setminus \{(1,2)\} = \{(1,3),(1,4),(1,5)\}$: If μ' is only allowed to be chosen u.a.r. from $S \setminus \{(1,2)\}$, then the probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;1}$ misses $(1,2)$ becomes $\frac{1}{3} \times \frac{1}{6} \times 15 = \frac{5}{6}$. Similarly, the probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;2}$ misses $(1,2)$ becomes $\frac{1}{3} \times \frac{1}{36} \times 81 = \frac{3}{4}$, which equals the probability that $\mathcal{W}^1_{\mathcal{K}_5;2}$ misses $(1,2)$. One may think that this construction solves the problem of how the two probabilities relate, however, if we consider $\ell = \ell' = 3$, the probability that $\mathcal{W}^1_{\mathcal{K}_5;3}$ misses $(1,2)$ is $\frac{21}{32}$ and the probability that $\mathcal{W}^{\mu'}_{\mathcal{K}'_5;3}$ misses $(1,2)$ is $\frac{145}{216}$.

In either way of choosing μ and μ', we have shown that the probability that a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ of length ℓ on $\mathcal{G}$ visits e is generally *not* the same as the probability that a random walk $\mathcal{W}^{\mu'}_{\mathcal{G}';\ell'}$ of length ℓ' on $\mathcal{G}'$ visits v_e, where $\ell' = \ell - 1$ or $\ell' = \ell$. However, as mentioned earlier, there exists a mapping between the probabilities as there is a one-to-one correspondence between $\mathcal{G}$ and $\mathcal{G}'$. Deriving the mapping is not straightforward and the mapping might not be easily extended to the case where we consider a random walk missing a longer prime. Therefore, we look for more direct approaches to understand the probability of a random walk missing an edge. These methods are covered in the next three sections.

5 Coverage probability of an edge using the walk generating function

Let $\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}$ be the event that a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ visits a given edge $e \in \mathcal{E}(\mathcal{G})$. When $\mathcal{G} = \mathcal{G}_r$, Lemma 1 gives

$$\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right) = 1 - \frac{m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)}{m\left(\mathfrak{W}^{\mu}_{\mathcal{G};\ell}\right)}. \tag{3}$$

As we know $m\left(\mathfrak{W}^{\mu}_{\mathcal{G};\ell}\right) = r^{\ell}$ when $\mathcal{G}$ is an r-regular graph, it is sufficient to get $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ in order to obtain $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$. In Section 5.1, we use the walk generating function to obtain an explicit expression of $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$. In Section 5.2, we use this method to derive an explicit expression of $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ when $\mathcal{G}$ is a complete or circle graph, and the starting point μ of the random walk is one of the endpoints of e.

5.1 The walk generating function

On any graph $\mathcal{G}$ which satisfies (P1) to (P5) stated in Section 2, we can calculate the number of walks from one vertex to another and of arbitrary length through its walk generating function [4]. We define the walk generating function $g^{\mu\nu}_{\mathcal{G}}(z)$ associated with $\{m(\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell})\}_{\ell\geq 0}$ as

$$g^{\mu\nu}_{\mathcal{G}}(z) = \sum_{\ell=0}^{\infty} m(\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell}) z^{\ell},$$

where $z \in \mathbb{C}$; the walk generating function $g^{\mu\nu}_{\mathcal{G}\backslash\{e\}}(z)$ associated with $\{m(\mathfrak{W}^{\mu\nu}_{\mathcal{G}\backslash\{e\};\ell})\}_{\ell\geq 0}$ is defined similarly. The radius of convergence of this series, denoted by $R_{\mu\nu}$, shall be given a little later in terms of the adjacency matrix of $\mathcal{G}$ using Eq. (5). In the following, we assume that $|z| < R_{\mu\nu}$. Then the walk generating function $g^{\mu}_{\mathcal{G}}(z)$ associated with $\{m(\mathfrak{W}^{\mu}_{\mathcal{G};\ell})\}_{\ell\geq 0}$ is defined as

$$g^{\mu}_{\mathcal{G}}(z) = \sum_{\nu\in\mathcal{V}} g^{\mu\nu}_{\mathcal{G}}(z) = \sum_{\ell=0}^{\infty}\sum_{\nu\in\mathcal{V}} m(\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell})z^{\ell} = \sum_{\ell=0}^{\infty} m(\mathfrak{W}^{\mu}_{\mathcal{G};\ell})z^{\ell}, \tag{4}$$

and the walk generating function $g^{\mu}_{\mathcal{G}\backslash\{e\}}(z)$ associated with $\{m(\mathfrak{W}^{\mu}_{\mathcal{G}\backslash\{e\};\ell})\}_{\ell\geq 0}$ is defined similarly.

We denote the walk generating matrix as $G_{\mathcal{G}}(z) = \left(g^{\mu\nu}_{\mathcal{G}}(z)\right)_{\mu,\nu\in\mathcal{V}(\mathcal{G})}$, where $|z| < \min_{\mu,\nu\in\mathcal{V}(\mathcal{G})} R_{\mu\nu}$. Let $A_{\mathcal{G}}$ be the adjacency matrix of $\mathcal{G}$. Notice that $m(\mathfrak{W}^{\mu\nu}_{\mathcal{G};\ell})$ is the (μ,ν) entry of $A^{\ell}_{\mathcal{G}}$. We can write

$$G_{\mathcal{G}}(z) = \sum_{\ell=0}^{\infty}(zA_{\mathcal{G}})^{\ell}. \tag{5}$$

Below are some properties of $G_{\mathcal{G}}(z)$.

(a) Radius of convergence: The series converges when $||zA_{\mathcal{G}}|| = |z|||A_{\mathcal{G}}|| < 1$ (see Theorem 7.3.1, [26] page 375). Here $||\cdot||$ is the spectral norm. Since $A_{\mathcal{G}}$ is not a zero matrix, and each entry of $A_{\mathcal{G}}$ is 0 or 1, we have $0 < ||A_{\mathcal{G}}|| < \infty$. The region of convergence of the series in Eq. (5) is $\{z : |z| < \frac{1}{||A_{\mathcal{G}}||}\}$. We define $G_{\mathcal{G}}(0)$ to be an identity matrix I_n of size n. We shall write $I_n = I$ when it is clear in the context what n is. When $|z| < \frac{1}{||A_{\mathcal{G}}||}$, we can write $G_{\mathcal{G}}(z) = (I_n - zA_{\mathcal{G}})^{-1}$ (Theorem 7.3.1, [26] page 375).

(b) Relationship with resolvent: When $0 < |z| < \frac{1}{||A_\mathcal{G}||}$, we can also write

$$\begin{aligned} G_\mathcal{G}(z) &= (I - zA_\mathcal{G})^{-1} \\ &= \frac{1}{z}\left(\frac{1}{z}I - A_\mathcal{G}\right)^{-1} \\ &= -\frac{1}{z}\left(A_\mathcal{G} - \frac{1}{z}I\right)^{-1}, \end{aligned} \tag{6}$$

where $\left(A_\mathcal{G} - \frac{1}{z}I\right)^{-1}$ is the resolvent of $A_\mathcal{G}$ [26]. We shall use some nice properties of the resolvent to illustrate that each entry of $(I - zA_\mathcal{G})^{-1}$ is holomorphic except when $\frac{1}{z}$ is an eigenvalue of $A_\mathcal{G}$. Let $z_1 = \frac{1}{z}$. Here we only list the properties we need and the reader is referred to [26] for the proofs of these properties (except for the first property listed below, which is immediate from Eq. (6)).

1. $(A_\mathcal{G} - z_1 I)^{-1}$ is undefined when z_1 is an eigenvalue of $A_\mathcal{G}$.

2. Eigenvalues of $A_\mathcal{G}$ lie in the disk given by $\{z_1 : |z_1| \leq ||A_\mathcal{G}||\}$ (Theorem 7.3.4, [26] page 377).

3. Each entry of $(A_\mathcal{G} - z_1 I)^{-1}$ is holomorphic at every point $z_1 \in \mathbb{C} \setminus \{\lambda_1, \ldots, \lambda_n\}$, where $\{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues of $A_\mathcal{G}$ (Theorem 7.5.2, [26] page 389).

Remark 1. $G_\mathcal{G}(z) = \sum_{\ell=0}^{\infty}(zA_\mathcal{G})^\ell$ coincides with $(I - zA_\mathcal{G})^{-1} = -\frac{1}{z}\left(A_\mathcal{G} - \frac{1}{z}I\right)^{-1}$ when $0 < |z| < \frac{1}{||A_\mathcal{G}||}$. However, $\left(A_\mathcal{G} - \frac{1}{z}I\right)^{-1}$ exists at every point $\frac{1}{z} \in \mathbb{C} \setminus \{\lambda_1, \ldots, \lambda_n\}$, and by Property 3, each entry of $(I - zA_\mathcal{G})^{-1}$ is holomorphic except when $\frac{1}{z} \in \{\lambda_1, \ldots, \lambda_n\}$.

Step 1: Obtain the walk generating functions

From the previous part, we know that $G_\mathcal{G}(z) = (I - zA_\mathcal{G})^{-1}$ and $(I - zA_\mathcal{G})^{-1}$ exists when $|z| < \frac{1}{||A_\mathcal{G}||}$. In special cases such as when $\mathcal{G} = \mathcal{K}_n$, we can obtain $(I - zA_\mathcal{G})^{-1}$ by using the blockwise inversion of $I - zA_{\mathcal{G}}$. For details, we refer the reader to Appendix A and Appendix B.1.

Step 2: Obtain the number of walks

Once we have calculated generating functions such as $g_{\mathcal{G}}^{\mu\nu}(z)$, we may recover $m(\mathfrak{W}_{\mathcal{G};\ell}^{\mu\nu})$ through the ℓ-th derivative of the walk generating functions. However, this can be a laborious task and we use instead a particular integral transform, the inverse $\mathcal{Z}$-transform.

The $\mathcal{Z}$-transform and its inverse connect a series $\{m(\mathfrak{W}_{\mathcal{G};\ell}^{\mu\nu})\}_{\ell\geq 0}$ with its ordinary generating function $g_{\mathcal{G}}^{\mu\nu}$ and back [23]. Assuming that $|z| > ||A_{\mathcal{G}}||$,

$$\mathcal{Z}[\{m(\mathfrak{W}_{\mathcal{G};\ell}^{\mu\nu})\}_{\ell\geq 0}](z) = \sum_{\ell=0}^{\infty} m(\mathfrak{W}_{\mathcal{G};\ell}^{\mu\nu}) z^{-\ell} = g_{\mathcal{G}}^{\mu\nu}(1/z),$$
$$\mathcal{Z}^{-1}[g_{\mathcal{G}}^{\mu\nu}(1/z)](\ell) = m(\mathfrak{W}_{\mathcal{G};\ell}^{\mu\nu}).$$

We know that $G_{\mathcal{G}}(1/z) = \sum_{\ell=0}^{\infty} \left(\frac{1}{z}A_{\mathcal{G}}\right)^{\ell}$. The region of convergence of the series is $\{z : |z| > ||A_{\mathcal{G}}||\}$. The inverse $\mathcal{Z}$-transform is given by a contour integral [23]

$$\mathcal{Z}^{-1}[g_{\mathcal{G}}^{\mu\nu}(1/z)](\ell) = \frac{1}{2\pi i} \oint_C g_{\mathcal{G}}^{\mu\nu}(1/z) z^{\ell-1} dz, \tag{7}$$

with C a contour on which $|z| > ||A_{\mathcal{G}}||$.

Following Remark 1, $G_{\mathcal{G}}(1/z) = \sum_{\ell=0}^{\infty} \left(\frac{1}{z}A_{\mathcal{G}}\right)^{\ell}$ coincides with $(I - \frac{1}{z}A_{\mathcal{G}})^{-1} = -z(A_{\mathcal{G}} - z)^{-1}$ when $|z| > ||A_{\mathcal{G}}||$. Each entry of $(I - \frac{1}{z}A_{\mathcal{G}})^{-1}$ is holomorphic except when $z \in \{\lambda_1, \dots, \lambda_n\}$, which means for any pairs of $\mu, \nu \in \mathcal{V}$, $g_{\mathcal{G}}^{\mu\nu}(1/z)$ is holomorphic except when $z \in \{\lambda_1, \dots, \lambda_n\}$. By the Cauchy Residue Theorem (Theorem 18.3, [34] page 211), the integral in Eq. (7) is equivalent to a sum of analytically available complex numbers,

$$\mathcal{Z}^{-1}[g^{\mu\nu}_{\mathcal{G}}(1/z)](\ell) = \sum_{z_0} \mathrm{Res}_{z_0}[g^{\mu\nu}_{\mathcal{G}}(1/z)]z^{\ell-1}, \tag{8}$$

$$\mathrm{Res}_{z_0}[g^{\mu\nu}_{\mathcal{G}}(1/z)]z^{\ell-1} = \frac{1}{(k-1)!}\lim_{z\to z_0}\frac{d^{k-1}}{dz^{k-1}}\left[(z-z_0)^k g^{\mu\nu}_{\mathcal{G}}(1/z)z^{\ell-1}\right], \tag{9}$$

with z_0 being a pole of $g^{\mu\nu}_{\mathcal{G}}(1/z)$ and k being the multiplicity of the pole.

Step 3: Obtain the probability

Using the walk generating function $g^{\mu\nu}_{\mathcal{G}\setminus\{e\}}(1/z)$ and the inverse $\mathcal{Z}$-transform, we can obtain the number of walks $\mathcal{Z}^{-1}[g^{\mu\nu}_{\mathcal{G}\setminus\{e\}}(1/z)](\ell)$ of length ℓ from μ to ν. Then we obtain the number of walks $\mathcal{Z}^{-1}[g^{\mu}_{\mathcal{G}\setminus\{e\}}(1/z)](\ell)$ of length ℓ starting at μ by Eq. (4) and the linearity of the inverse $\mathcal{Z}$-transform [23]. Finally we get an explicit expression of $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e})$ by Eq. (3). We shall see the explicit expression of $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e})$ when $\mathcal{G}$ is a complete graph or a circle graph in the following section.

5.2 On a complete graph and a circle graph

Let $\mathcal{G}$ be a complete graph or a circle graph, and let $e = (\nu_1, \nu_2) \in \mathcal{E}(\mathcal{G})$. Note that for $\ell \geq 1$, by symmetry of $\mathcal{G}$, $\mathbb{P}^{\nu_1}_{\mathcal{G};\ell}\left(\mathcal{A}^{\nu_1}_{\mathcal{G};\ell;e}\right) = \mathbb{P}^{\nu_2}_{\mathcal{G};\ell}\left(\mathcal{A}^{\nu_2}_{\mathcal{G};\ell;e}\right)$, i.e. we obtain the same probability regardless of which endpoint of e is chosen as the starting point of a random walk. Without loss of generality, assume that $\nu_1 \equiv 1$ and $\nu_2 \equiv 2$. Let $\{1, 2, \ldots, n\}$ be the labels of the vertices in $\mathcal{G}$, labeled clockwise. If a random walk starts from 2 instead of 1, one can relabel the vertices anti-clockwise using $\{1, 2, \ldots, n\}$ (starting from the "original" vertex 2) without changing the probability of visiting e.

Theorem 3. *Let $e = (\nu_1, \nu_2) \in \mathcal{E}$, and $\mu = \nu_1$ or $\mu = \nu_2$. On a complete graph $\mathcal{K}_n$ with $n \geq 3$ vertices and for $\ell \geq 1$, we have*

$$\mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e}) = 1 - \frac{[n-1+c(n)][n-3+c(n)]^{\ell} - [n-1-c(n)][n-3-c(n)]^{\ell}}{(n-1)^{\ell}2^{\ell+1}c(n)},$$

where $c(n) = \sqrt{n^2+2n-7}$.

Proof. See Appendix B.1 for a proof using walk generating functions and Appendix B.2 for a proof using recurrence relations. □

Remark 2. To make sense of $\mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e}) = 1 - \frac{[n-1+c(n)][n-3+c(n)]^{\ell} - [n-1-c(n)][n-3-c(n)]^{\ell}}{(n-1)^{\ell}2^{\ell+1}c(n)}$, where $c(n) = \sqrt{n^2+2n-7}$, we write $c(n) = \sqrt{(n+1)^2 - 8} = \sqrt{(n+\gamma)^2}$ for $n \geq 4$, where $0 < \gamma < 1$. Then we see that, for $n \geq 4$,

$$\begin{aligned}
\mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e}) &= 1 - \frac{[n-1+c(n)][n-3+c(n)]^{\ell} - [n-1-c(n)][n-3-c(n)]^{\ell}}{(n-1)^{\ell}2^{\ell+1}c(n)} \\
&= 1 - \frac{(n-1+n+\gamma)(n-3+n+\gamma)^{\ell} - (n-1-n-\gamma)(n-3-n-\gamma)^{\ell}}{(n-1)^{\ell}2^{\ell+1}(n+\gamma)} \\
&= 1 - \frac{(2n-1+\gamma)(2n-3+\gamma)^{\ell} + (1+\gamma)(-1)^{\ell}(3+\gamma)^{\ell}}{(n-1)^{\ell}2^{\ell+1}(n+\gamma)} \\
&= 1 - \left(\frac{2n+\gamma-1}{2n+2\gamma}\right)\left(\frac{2n+\gamma-3}{2n-2}\right)^{\ell}\left[1 - \frac{(-1)^{\ell}}{\left(1+\frac{2n-2}{1+\gamma}\right)\left(1+\frac{2n-6}{3+\gamma}\right)^{\ell}}\right] \\
&= 1 - \left(1 - \frac{\gamma+1}{2n+2\gamma}\right)\left(1 - \frac{1-\gamma}{2n-2}\right)^{\ell}\left[1 - \frac{(-1)^{\ell}}{\left(1+\frac{2n-2}{1+\gamma}\right)\left(1+\frac{2n-6}{3+\gamma}\right)^{\ell}}\right].
\end{aligned}$$

Note that $1-\gamma > 0$ and

$$\begin{aligned}
1-\gamma &= (n+1) - \sqrt{(n+1)^2 - 8} \\
&= (n+1)\left[1 - \sqrt{1 - \frac{8}{(n+1)^2}}\right]
\end{aligned}$$

$$
\begin{aligned}
&= (n+1)\left[1 - 1 + \frac{1}{2}\frac{8}{(n+1)^2} + O\left(\frac{1}{n^4}\right)\right] \\
&= O\left(\frac{1}{n}\right),
\end{aligned}
$$

where the third step follows from the binomial expansion of $\left[1-\frac{8}{(n+1)^2}\right]^{\frac{1}{2}}$. Then $1+\gamma = 2 - O\left(\frac{1}{n}\right)$ and $3+\gamma = 4 - O\left(\frac{1}{n}\right)$.Thus we can see that,

(a) When n is fixed and $\ell \to \infty$: $1 - \mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e})$ decays exponentially in ℓ.

(b) When ℓ is fixed and $n \to \infty$: $\mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e})$ decays polynomially in n.

Theorem 4. *Let $e = (\nu_1, \nu_2) \in \mathcal{E}$, and $\mu = \nu_1$ or $\mu = \nu_2$. On a circle graph $\mathcal{C}_n$ with $n \geq 3$ vertices and for $\ell \geq 1$, we have*

$$
\mathbb{P}^{\nu_1}_{\mathcal{C}_n;\ell}(\mathcal{A}^{\nu_1}_{\mathcal{C}_n;\ell;e}) = 1 - \sum_{d=0}^{n-1}\sum_{k=1}^{n} \frac{\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell} \prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi j}{n-d}\right)\right]}{2^d \prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi h}{n+1}\right)\right]}.
$$

Proof. See Appendix B.3. □

Remark 3. To make sense of $\mathbb{P}^{\nu_1}_{\mathcal{C}_n;\ell}(\mathcal{A}^{\nu_1}_{\mathcal{C}_n;\ell;e})$, we see that

$$
\begin{aligned}
&\mathbb{P}^{\mu}_{\mathcal{C}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;e}) \\
&= 1 - \sum_{d=0}^{n-1}\sum_{k=1}^{n} \frac{\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell} \prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi j}{n-d}\right)\right]}{2^d \prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi h}{n+1}\right)\right]} \\
&= 1 - \sum_{k=1}^{n}\left\{\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell} \sum_{d=0}^{n-1} \frac{\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi j}{n-d}\right)\right]}{2^d \prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi h}{n+1}\right)\right]}\right\} \\
&= 1 - \sum_{k=1}^{n}\left\{\left[1 - \frac{1}{2}\left(\frac{\pi k}{n+1}\right)^2 + O\left(\frac{1}{n^4}\right)\right]^{\ell} \sum_{d=0}^{n-1} \frac{\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi j}{n-d}\right)\right]}{2^d \prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right) - \cos\left(\frac{\pi h}{n+1}\right)\right]}\right\},
\end{aligned}
$$

where we use the Taylor expansion of $\cos\left(\frac{\pi k}{n+1}\right)$ in the last step.

(a) When n is fixed and $\ell \to \infty$: $1 - \mathbb{P}^{\mu}_{\mathcal{C}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;e})$ decays exponentially in ℓ.

(b) When ℓ is fixed: When $\ell \leq n-1$ and $n \leq n'$, $\mathbb{P}^{\mu}_{\mathcal{C}_n;\ell}(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;e}) = \mathbb{P}^{\mu}_{\mathcal{C}_{n'};\ell}(\mathcal{A}^{\mu}_{\mathcal{C}_{n'};\ell;e})$ as the number of walks, starting from ν_1 or ν_2, of length ℓ on $\mathcal{C}_n$ that do not visit e equals the number of walks, starting from ν_1 or ν_2, of length ℓ on $\mathcal{C}_{n'}$ that do not visit e.

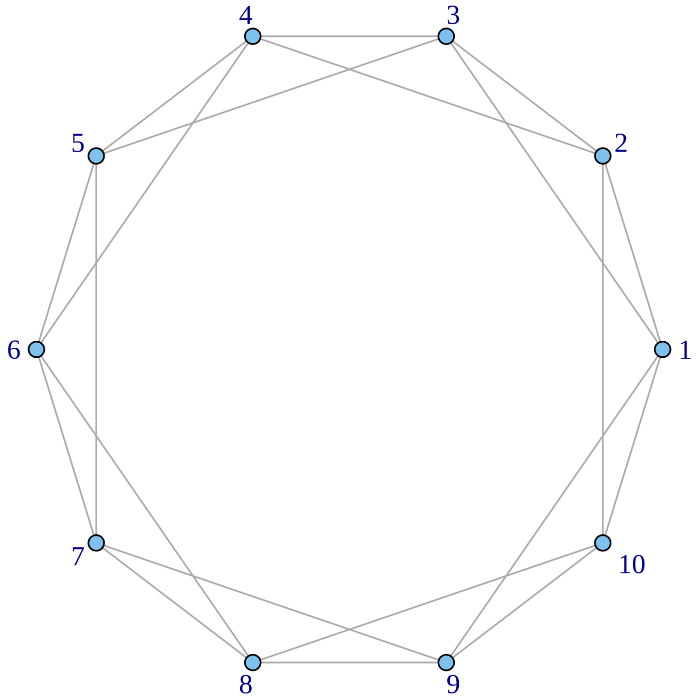


Figure 3: The 2-nearest neighbour graph on 10 vertices, $\mathcal{G}(10,2)$ (unique up to isomorphism).

When $2 < r < n-1$, the expressions of the walk generating functions $g^{\mu\nu}_{\mathcal{G}_r}(z)$ depend on the structure of $\mathcal{G}_r$. We focus on a specific type of regular graph constructed as follows: i) construct a circle graph on n vertices; and ii) place edges between all k nearest neighbours on the circle graph, where $1 \le k \le \lfloor \frac{n}{2} \rfloor$. We call it a *$k$-nearest-neighbour graph* and denote it by $\mathcal{G}(n,k)$. The 2-nearest neighbour graph on 10 vertices $\mathcal{G}(10,2)$ (unique up to isomorphism) is shown in Fig. 3. However, the expressions of the walk generating functions are very involved and lack elegance, and thus we do not give them here. We could not use the inverse $\mathcal{Z}$-transform to obtain $m\left(\mathfrak{W}^{\mu\nu}_{\mathcal{G}(n,k)\setminus\{e\};\ell}\right)$, where $e \in \mathcal{E}(\mathcal{G}(n,k))$. In the next section, we use the largest eigenvalue of $\mathcal{G}_r \setminus \{e\}$ to obtain an upper bound for $m\left(\mathfrak{W}^{\mu\nu}_{\mathcal{G}_r\setminus\{e\};\ell}\right)$.

6 Coverage probability of an edge using the largest eigenvalue of the graph

We can provide an upper bound for $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ by using the largest eigenvalue of $A_{\mathcal{G}\setminus\{e\}}$, which in turns gives a lower bound for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ through Eq. (3). As $A_{\mathcal{G}\setminus\{e\}}$ is a symmetric real matrix, its eigenvalues are real. There are more properties of $A_{\mathcal{G}\setminus\{e\}}$ which we can use to say something about its largest eigenvalue. We denote the largest eigenvalue of a matrix A by $\lambda_{max}(A)$.

Step 0: Preliminaries

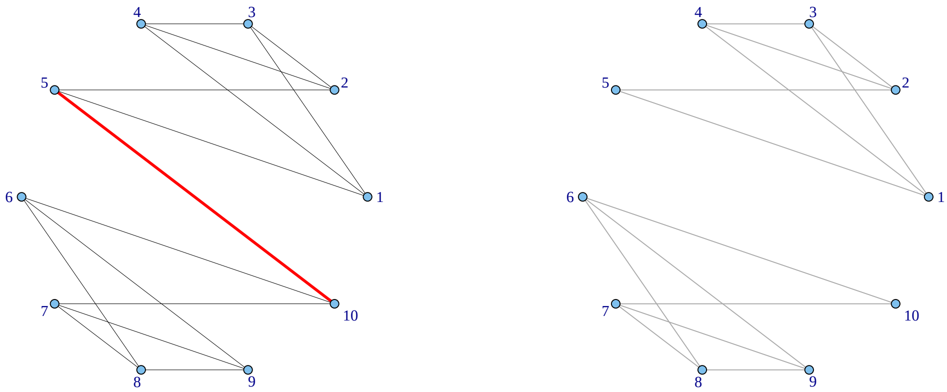
Let $\mathcal{M}_n$ denote the set of $n \times n$ real matrices.

Definition 1. ([22](Definition 6.2.21)) *A matrix $A \in \mathcal{M}_n$ is said to be reducible if $n = 1$ and $A = 0$, or for $n \geq 2$, there is a permutation matrix $P \in \mathcal{M}_n$ such that*

$$P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

where $B \in \mathcal{M}_r$ and $D \in \mathcal{M}_{n-r}$, for some r where $1 \leq r \leq n-1$. A is said to be irreducible if it is not reducible.

A permutation matrix $P \in \mathcal{M}_n$ is a square matrix with exactly one entry in each row and column being equal to 1 and all other entries being 0. Thus $P^T \in \mathcal{M}_n$ is also a permutation matrix. Left multiplication of a matrix $A \in \mathcal{M}_n$ by a permutation matrix $P \in \mathcal{M}_n$ permutes the rows of A, while right multiplication of A by P permutes the columns of A. We recall that $\mathcal{G}$ is connected (see Section 2). $\mathcal{G}$ is connected if and only if $A_{\mathcal{G}}$ is irreducible [7].



(a) (b)

Figure 4: (a) A 3-regular graph $\mathcal{G}_3$ with a cut edge (5,10); (b) Removing (5,10) from $\mathcal{G}_3$ gives a graph with two connected components.

If $\mathcal{G}$ is connected, $\mathcal{G} \setminus \{e\}$ is connected, where $e \in \mathcal{E}(\mathcal{G})$, unless e is a cut edge (or bridge). Fig. 4a shows a 3-regular graph $\mathcal{G}_3$ with a cut edge (5,10). When e is a cut edge, removing e from $\mathcal{G}$ gives a graph $\mathcal{G} \setminus \{e\}$ with two connected components, say $\mathcal{G}^1$ and $\mathcal{G}^2$. However the adjacency matrices $A_{\mathcal{G}^1}$ and $A_{\mathcal{G}^2}$ of the connected components $\mathcal{G}^1$ and $\mathcal{G}^2$ respectively are irreducible (for example see Fig. 4b). We shall see why this observation is important a little later.

Definition 2. ([22](Definition 1.2.9)) *Let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be the (not necessarily distinct) eigenvalues of* $A \in \mathcal{M}_n$. *The spectral radius of* A *is a nonnegative real number* $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$.

Theorem 5. (Perron-Frobenius Theorem, [22](Theorem 8.4.4)) *Let $A \in \mathcal{M}_n$ be irreducible and nonnegative. Then*

(a) $\rho(A) > 0$*;*

(b) $\rho(A)$ *is an eigenvalue of* A*, i.e.* $\rho(A)$ *is the largest eigenvalue of* A*;*

(c) There is a positive vector v *(all entries are positive), where* $v^T v = 1$*, such that* $Av = \rho(A)v$*; and*

(d) $\rho(A)$ *is a simple eigenvalue of* A*.*

If $\mathcal{G} \setminus \{e\}$ is connected, i.e. $A_{\mathcal{G}\setminus\{e\}}$ is irreducible, the Perron-Frobenius Theorem ensures that the largest eigenvalue of $A_{\mathcal{G}\setminus\{e\}}$ is positive and simple, and the associated normalised eigenvector is positive.

Step 1: Obtain an upper bound for the number of walks

We recall that $m\left(\mathfrak{W}^{\mu\nu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ is the (μ,ν) entry of of $A^{\ell}_{\mathcal{G}\setminus\{e\}}$. Consequently, $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) = \sum_{\nu\in\mathcal{V}(\mathcal{G}\setminus\{e\})} m\left(\mathfrak{W}^{\mu\nu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ is the μ-th row sum of $A^{\ell}_{\mathcal{G}\setminus\{e\}}$. The following lemma gives a link between row sums of a matrix and its spectral radius.

Lemma 6. ([22](Corollary 8.1.33)) *Let* $A \in \mathcal{M}_n$ *be nonnegative and write* $(A^{\ell})_{ij} = a^{(\ell)}_{ij}$*. If* A *has a positive eigenvector* $x = [x_i]$*, then for* $\ell \geq 1$ *and for all* $i = 1, \dots, n$ *we have*

$$\sum_{j=1}^{n} a^{(\ell)}_{ij} \leq \left(\frac{\max_{1\leq k\leq n} x_k}{\min_{1\leq k\leq n} x_k}\right) \rho(A)^{\ell}.$$

(a) When $e \in \mathcal{E}(\mathcal{G})$ is not a cut edge: If $A_{\mathcal{G}\setminus\{e\}} \in \mathcal{M}_n$ is irreducible, the spectral radius of $A_{\mathcal{G}\setminus\{e\}}$ is $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)$ by the Perron-Frobenius Theorem (Theorem 5). Thus we can provide an upper bound for $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ using Lemma 6.

Theorem 7. *Let $e \in \mathcal{E}(\mathcal{G})$ and $\mu \in \mathcal{V}(\mathcal{G})$. Assume that $\mathcal{G} \setminus \{e\}$ is connected. Let $x = [x_i]$ be the positive normalised eigenvector paired with $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)$. For $\ell \geq 1$, we have*

$$m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq c(x)\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)^{\ell},$$

where $c(x) = \left(\frac{\max_{1\leq k\leq n} x_k}{\min_{1\leq k\leq n} x_k}\right) \geq 1$ (note that $c(x)$ depends on n).

Proof. As $A_{\mathcal{G}\setminus\{e\}}$ is nonnegative and irreducible, by the Perron-Frobenius Theorem (Theorem 5), $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right) = \rho\left(A_{\mathcal{G}\setminus\{e\}}\right) > 0$ and there is a positive normalised eigenvector $x = [x_i]$ paired with $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)$. Since $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)$ is an eigenvalue of $A_{\mathcal{G}\setminus\{e\}}$ and $\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right) = \rho\left(A_{\mathcal{G}\setminus\{e\}}\right) > 0$, we have $\lambda_{max}\left(A^{\ell}_{\mathcal{G}\setminus\{e\}}\right) = \lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)^{\ell}$.

As $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ is the μ-th row sum of $A^{\ell}_{\mathcal{G}\setminus\{e\}}$, by Lemma 6, we have $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq c(x)\lambda_{max}\left(A_{\mathcal{G}\setminus\{e\}}\right)^{\ell}$, where $c(x) = \left(\frac{\max_{1\leq k\leq n} x_k}{\min_{1\leq k\leq n} x_k}\right) \geq 1$. □

(b) When $e \in \mathcal{E}(\mathcal{G})$ is a cut edge: If $e \in \mathcal{E}(\mathcal{G})$ is a cut edge, $A_{\mathcal{G}\setminus\{e\}} \in \mathcal{M}_n$ is reducible. As what we have mentioned earlier, $\mathcal{G} \setminus \{e\}$ has two connected components, say $\mathcal{G}^1$ and $\mathcal{G}^2$. Excluding the case where $\mathcal{G}^1$ or $\mathcal{G}^2$ is a singleton, we can still obtain a similar upper bound for $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ as in Theorem 7. If the starting point μ is one of the vertices in $\mathcal{G}^1$, the number of walks of length ℓ on $\mathcal{G} \setminus \{e\}$ that do not visit e equals the number of walks of length ℓ on $\mathcal{G}^1$, starting at μ (for example see Fig. 4b, where $\mu = 1$). The argument is similar if $\mu \in \mathcal{V}(\mathcal{G}^2)$.

Theorem 8. *Let $e \in \mathcal{E}(\mathcal{G})$. Assume that $\mathcal{G} \setminus \{e\}$ has two connected components $\mathcal{G}^1$ and $\mathcal{G}^2$ with $|\mathcal{V}(\mathcal{G}^1)| = n_1$ and $|\mathcal{V}(\mathcal{G}^2)| = n_2 = n - n_1$, where $1 < n_1 < n-1$. Let $x' = [x'_i]$ be the positive normalised eigenvector paired with $\lambda_{max}\left(A_{\mathcal{G}^1}\right)$, and $y = [y_i]$ be the positive* *normalised eigenvector paired with* $\lambda_{max}(A_{\mathcal{G}^2})$*. If* $\mu \in \mathcal{V}(\mathcal{G}^1)$*, for* $\ell \geq 1$*, we have*

$$m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq c(x')\lambda_{max}\left(A_{\mathcal{G}^1}\right)^{\ell},$$

where $c(x') = \left(\frac{\max_{1\leq k\leq n_1} x'_k}{\min_{1\leq k\leq n_1} x'_k}\right) \geq 1$ *(note that* $c(x)$ *depends on* n_1*). If* $\mu \in \mathcal{V}(\mathcal{G}^2)$*, for* $\ell \geq 1$*, we have*

$$m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq c(y)\lambda_{max}\left(A_{\mathcal{G}^2}\right)^{\ell},$$

where $c(y) = \left(\frac{\max_{1\leq k\leq n_2} y_k}{\min_{1\leq k\leq n_2} y_k}\right) \geq 1$ *(note that* $c(y)$ *depends on* n_2*).*

Proof. The proof follows directly from applying Theorem 7 to $\mathcal{G}^1$ and $\mathcal{G}^2$ separately. □

Step 3: Obtain an expression of the largest eigenvalue

When $\mathcal{G} = \mathcal{G}_r$, we can indeed obtain a lower bound for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e})$ by Eq. (3) if we can give an expression of the largest eigenvalue.

Theorem 9. ([28](Theorem 3.3)) *Let* G *be a connected non-regular graph on* n *vertices with maximum degree* d_{max}*. Then*

$$\lambda_{max}\left(A_{\mathcal{G}}\right) < d_{max} - \frac{d_{max}+1}{n(3n+d_{max}-8)}.$$

The upper bound for the largest eigenvalue of a connected non-regular graph shown in Theorem 9 (see [30] for an erratum to [28]) improved the results in [31] and [37]. Liu *et al.* gave a sharper upper bound for the largest eigenvalue of a connected non-regular graph in [29]. We were able to follow their argument but were unable to justify a step in the proof of the result. Hence we do not use the result here.

(a) When $e \in \mathcal{E}(\mathcal{G}_r)$ is not a cut edge: Notice that $\mathcal{G}_r \backslash \{e\}$ is a connected non-regular graph with maximum degree r. Thus

$$\lambda_{max}\left(A_{\mathcal{G}_r \backslash \{e\}}\right) < r - \frac{r+1}{n(3n+r-8)}. \tag{10}$$

(b) When $e \in \mathcal{E}(\mathcal{G}_r)$ is a cut edge: Similarly, upper bounds for $\lambda_{max}\left(A_{\mathcal{G}_r^1}\right)$ and $\lambda_{max}\left(A_{\mathcal{G}_r^2}\right)$, where $\mathcal{G}_r^1$ and $\mathcal{G}_r^2$ are two connected components of $\mathcal{G}_r$, are

$$\lambda_{max}\left(A_{\mathcal{G}_r^1}\right) < r - \frac{r+1}{n_1(3n_1+r-8)} \quad \text{and} \tag{11}$$

$$\lambda_{max}\left(A_{\mathcal{G}_r^2}\right) < r - \frac{r+1}{n_2(3n_2+r-8)} \tag{12}$$

respectively. Here $n_1 = |\mathcal{V}(\mathcal{G}_r^1)|$ and $n_2 = |\mathcal{V}(\mathcal{G}_r^2)| = n - n_1$.

Remark 4. When $r = 2$, the only connected $\mathcal{G}_2$ is $\mathcal{C}_n$ and $\mathcal{C}_n \setminus \{e\}$ is a path graph. It is well known that

$$\begin{aligned} \lambda_{max}\left(A_{\mathcal{C}_n \backslash \{e\}}\right) &= 2\cos\left(\frac{\pi}{n+1}\right) \\ &= 2 - \frac{\pi^2}{(n+1)^2} + O\left(\frac{1}{n^4}\right), \end{aligned}$$

where the last step follows from taking the Taylor expansion of $\cos\left(\frac{\pi}{n+1}\right)$. Using Inequality (10), one can get $\lambda_{max}\left(A_{\mathcal{C}_n \backslash \{e\}}\right) < 2 - \frac{1}{n(n-2)}$. Thus for large n, the upper bound for $\lambda_{max}\left(A_{\mathcal{C}_n \backslash \{e\}}\right)$ given in Inequality (10) is reasonably sharp. When $r = n-1$,

$\mathcal{G}_r = \mathcal{K}_n$ and $\lambda_{max}\left(A_{\mathcal{K}_n \setminus \{e\}}\right) = \frac{n-3+\sqrt{n^2+2n-7}}{2}$ [36](Theorem 2.2). Note that

$$\begin{aligned}
\frac{n-3+\sqrt{n^2+2n-7}}{2} &= n-1-\frac{(n+1)\left[1-\sqrt{1-\frac{8}{(n+1)^2}}\right]}{2} \\
&= n-1-\frac{(n+1)\left[1-1+\frac{4}{(n+1)^2}+O\left(\frac{1}{n^4}\right)\right]}{2} \\
&= n-1-\frac{2}{n+1}-O\left(\frac{1}{n^4}\right),
\end{aligned}$$

where the second last step follows from taking the Taylor expansion of $\left[1-\frac{8}{(n+1)^2}\right]^{\frac{1}{2}}$. Again, using Inequality (10), one can get $\lambda_{max}\left(A_{\mathcal{K}_n \setminus \{e\}}\right) < n-1-\frac{1}{4n-9}$. Thus for large n, the upper bound for $\lambda_{max}\left(A_{\mathcal{K}_n \setminus \{e\}}\right)$ given in Inequality (10) is reasonably sharp.

Remark 5. If one uses the Frobenius Theorem [22](Theorem 8.1.22)), it is straightforward that $r-1 \leq \lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right) \leq r$, where $2 \leq r \leq n-1$. Assume that $\mathcal{G}_r$ and $\mathcal{G}_r \setminus \{e\}$ are connected. If one uses the fact that $\lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right)$ is at least the average degree of $\mathcal{G}_r \setminus \{e\}$ [7](Property 2.1.E2) and $\lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right) < \lambda_{max}\left(A_{\mathcal{G}_r}\right)$ (see Appendix C for a proof), then it is possible to obtain

$$\lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right) = r - \epsilon(n),$$

where $0 < \epsilon(n) \leq \frac{2}{n}$. However it is obvious that the upper bound $\lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right) < r$ is not sharper than the upper bound given in Inequality (10).

Step 4: Obtain a lower bound for the probability

Having derived an upper bound for $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r \setminus \{e\};\ell}\right)$, we are now in a position to obtain a lower bound for $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e})$.

Theorem 10. *Let $e \in \mathcal{E}(\mathcal{G}_r)$ and $\mu \in \mathcal{V}(\mathcal{G}_r)$. Assume that $\mathcal{G}_r \setminus \{e\}$ is connected. Let $x = [x_i]$ be the positive eigenvector paired with $\lambda_{max}\left(A_{\mathcal{G}_r \setminus \{e\}}\right)$. On a fixed r-regular graph $\mathcal{G}_r$, where $2 < r < n-1$, and for $\ell \geq 1$, we have*

$$\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e}) > 1 - c(x)\left[1 - \frac{r+1}{nr(3n+r-8)}\right]^{\ell}$$

where $c(x) = \left(\frac{\max_{1\leq k\leq n} x_k}{\min_{1\leq k\leq n} x_k}\right) \geq 1$ (note that $c(x)$ depends on n).

Proof. This is immediate by using Theorem 7, Eq. (10) and Eq. (3) with $\mathcal{G} = \mathcal{G}_r$ and $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}\right) = r^{\ell}$. □

Theorem 11. *Let $e \in \mathcal{E}(\mathcal{G}_r)$ and $2 < r < n-1$. Assume that $\mathcal{G}_r \setminus \{e\}$ has two connected components $\mathcal{G}^1_r$ and $\mathcal{G}^2_r$ with $|\mathcal{V}(\mathcal{G}^1_r)| = n_1$ and $|\mathcal{V}(\mathcal{G}^2_r)| = n_2 = n - n_1$, where $1 < n_1 < n-1$. Let $x' = [x'_i]$ be the positive normalised eigenvector paired with $\lambda_{max}\left(A_{\mathcal{G}^1_r}\right)$, and $y = [y_i]$ be the positive normalised eigenvector paired with $\lambda_{max}\left(A_{\mathcal{G}^2_r}\right)$. If $\mu \in \mathcal{V}(\mathcal{G}^1_r)$, for $\ell \geq 1$, we have*

$$\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e}) > 1 - c(x')\left[1 - \frac{r+1}{n_1 r(3n_1+r-8)}\right]^{\ell},$$

where $c(x') = \left(\frac{\max_{1\leq k\leq n_1} x'_k}{\min_{1\leq k\leq n_1} x'_k}\right) \geq 1$ (note that $c(x')$ depends on n_1). If $\mu \in \mathcal{V}(\mathcal{G}^2_r)$, for $\ell \geq 1$, we have

$$\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e}) > 1 - c(y)\left[1 - \frac{r+1}{n_2 r(3n_2+r-8)}\right]^{\ell},$$

where $c(y) = \left(\frac{\max_{1\leq k\leq n_2} y_k}{\min_{1\leq k\leq n_2} y_k}\right) \geq 1$ (note that $c(y)$ depends on n_2).

Proof. This is immediate by using Theorem 8, Eq. (11) to $\mathcal{G}^1_r$, Eq. (12) to $\mathcal{G}^2_r$ and Eq. (3) with $\mathcal{G} = \mathcal{G}_r$ and $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}_r;\ell}\right) = r^{\ell}$. □

Remark 6.

(a) From Theorem 10 and Theorem 11, we can see that for $2 < r < n-1$, when n is fixed, $1 - \mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e})$ decays exponentially in ℓ. Notice that there is no cut edge in a complete graph or in a circle graph. Let $x = [x_i]$ and $y = [y_i]$ be the positive eigenvectors paired with $\lambda_{max}\left(A_{\mathcal{C}_n\setminus\{e\}}\right)$ and $\lambda_{max}\left(A_{\mathcal{K}_n\setminus\{e\}}\right)$ respectively. Following Remark 4, we have

$$\mathbb{P}^{\mu}_{\mathcal{C}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;e}\right) \geq 1 - c(x)\left[\cos\left(\frac{\pi}{n+1}\right)\right]^{\ell},$$

where $c(x) = \left(\frac{\max_{1\leq k\leq n} x_k}{\min_{1\leq k\leq n} x_k}\right) \geq 1$, and

$$\begin{aligned}\mathbb{P}^{\mu}_{\mathcal{K}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e}\right) &\geq 1 - c(y)\left[\frac{n-3+\sqrt{n^2+2n-7}}{2(n-1)}\right]^{\ell} \\ &= 1 - c(y)\left(1 - \frac{1-\gamma}{2n-2}\right)^{\ell},\end{aligned}$$

where $c(y) = \left(\frac{\max_{1\leq k\leq n} y_k}{\min_{1\leq k\leq n} y_k}\right) \geq 1$. We recall that $\sqrt{n^2+2n-7} = n+\gamma$, where $0 < \gamma < 1$, and $1-\gamma = O\left(\frac{1}{n}\right)$. Thus exponential decay of $1 - \mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e})$ when $\mathcal{G}_r$ is a circle or complete graph is in agreement with the remarks following Theorem 3 and Theorem 4.

(b) Let $A \in \mathcal{M}_n$ be nonnegative. The Frobenius Theorem [22] (Theorem 8.1.22) gives

$$\min_i \sum_{j=1}^{n} a_{ij} \leq \rho(A) \leq \max_i \sum_{j=1}^{n} a_{ij}. \tag{13}$$

We recall that $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ is the μ-th row sum of $A^{\ell}_{\mathcal{G}\setminus\{e\}}$. Then we can write Inequality (13) as

$$\min_{\mu'\in\mathcal{V}} m\left(\mathfrak{W}^{\mu'}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq \lambda_{max}(A')^{\ell} \leq \max_{\mu'\in\mathcal{V}} m\left(\mathfrak{W}^{\mu'}_{\mathcal{G}\setminus\{e\};\ell}\right).$$

Here $A' = A_{\mathcal{G}\setminus\{e\}}$ if $\mathcal{G} \setminus \{e\}$ is connected. If $\mathcal{G} \setminus \{e\}$ has two connected components $\mathcal{G}^1$ and $\mathcal{G}^2$, then $A' = A_{\mathcal{G}^1}$ if $\mu \in \mathcal{V}(\mathcal{G}^1)$, or $A' = A_{\mathcal{G}^2}$ if $\mu \in \mathcal{V}(\mathcal{G}^2)$. There exists $\mu \in \mathcal{V}$ such that

$$m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right) = \min_{\mu' \in \mathcal{V}} m\left(\mathfrak{W}^{\mu'}_{\mathcal{G}\setminus\{e\};\ell}\right) \leq \lambda_{max}(A')^{\ell}.$$

which gives a sharper bound for $m\left(\mathfrak{W}^{\mu}_{\mathcal{G}\setminus\{e\};\ell}\right)$ than the one in Theorem 7 (unless $c(x) = 1$) or Theorem 8 (unless $c(x') = 1$ or $c(y) = 1$). This in turns give a sharper lower bound for $\mathbb{P}^{\mu}_{\mathcal{G}_r;\ell}(\mathcal{A}^{\mu}_{\mathcal{G}_r;\ell;e})$ than that in Theorem 10 (unless $c(x) = 1$) or Theorem 11 (unless $c(x') = 1$ or $c(y) = 1$).

7 Coverage probability of an edge on random regular graphs

In this section, we consider the case when $\mathcal{G}$ is chosen u.a.r. from a set of r-regular graphs on n vertices. In general, for a random graph on n vertices, the existence of an edge between two vertices is random. Let $\mathcal{V}$ be a given set of vertices, where $|\mathcal{V}| = n$. We let $\mathfrak{G}_n = \{\mathcal{G} = (\mathcal{V}, \mathcal{E}) : \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}\}$ be the set of graphs with vertex set $\mathcal{V}$ and $\mathcal{F}_{\mathfrak{G}_n}$ be the power set of $\mathfrak{G}_n$. We define a probability space of graphs by putting probabilities on the sets of edges, i.e. $\mathbb{P}_{\mathfrak{G}_n}(\mathbf{G} = \mathcal{G}) = \mathbb{P}_{\mathfrak{G}_n}(\mathbf{E} = \mathcal{E})$ is given by some law, where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$.

For any r-regular graph $\mathcal{G}$ on n vertices, we can define a probability measure $\mathbb{P}_{\mathcal{U}}$ for the walks on $\mathcal{G}$ by augmenting the set of r-regular graphs on n vertices with the set of all possible walks of length ℓ, starting from μ, on n vertices. For a fixed vertex set $\mathcal{V}$, where $|\mathcal{V}| = n$, let $\mathfrak{G}^{(r)}_n \subseteq \mathfrak{G}_n$ denote the set of r-regular graphs on n vertices and let $\mathcal{F}_{\mathfrak{G}_n^{(r)}}$ be the power set of $\mathfrak{G}_n^{(r)}$. Note that there are no isolated vertices in an r-regular graph. We recall that $\mathfrak{W}^{\mu}_{\mathcal{K}_n;\ell}$ is the set of all possible walks of length ℓ on $\mathcal{V}$, starting from μ, and let $\mathcal{F}_{\mathfrak{W}^{\mu}_{\mathcal{K}_n;\ell}}$ be the power set of $\mathfrak{W}^{\mu}_{\mathcal{K}_n;\ell}$. Augmenting $\mathfrak{G}_n^{(r)}$ with $\mathfrak{W}^{\mu}_{\mathcal{K}_n;\ell}$ by the Cartesian product of sets, we get $\mathcal{U} = \mathfrak{G}_n^{(r)} \times \mathfrak{W}^{\mu}_{\mathcal{K}_n;\ell}$ and a measurable space $(\mathcal{U}, \mathcal{F}_{\mathcal{U}})$, where $\mathcal{F}_{\mathcal{U}}$ is the power set of $\mathcal{U}$. We define a probability measure $\mathbb{P}_{\mathcal{U}}$ on the measurable space $(\mathcal{U}, \mathcal{F}_{\mathcal{U}})$ by taking $\mathbb{P}_{\mathcal{U}}(u) = \mathbb{P}_{\mathfrak{G}_n^{(r)}}(\mathcal{G})\mathbb{P}^{\mu}_{\mathcal{G};\ell}(w)$, where $u = (\mathcal{G}, w) \in \mathcal{F}_{\mathcal{U}}$, and $\mathbb{P}_{\mathfrak{G}_n^{(r)}}(\mathcal{G}_r^{rnd} = \mathcal{G}) = \frac{1}{\left|\mathfrak{G}_n^{(r)}\right|}$ is the uniform measure on $\mathfrak{G}_n^{(r)}$. An element $\mathcal{G}_r^{rnd} \in \mathfrak{G}_n^{(r)}$ chosen u.a.r. from $\mathfrak{G}_{n,r}$ is called a *random r-regular graph* [6]. Here $\mathbb{P}^{\mu}_{\mathcal{G};\ell}(w)$ is the probability measure on the measurable space of walks of length ℓ on $\mathcal{G}$ starting from μ, which is given by $(*)$ (see Section 3). Thus we now have two layers of randomness: not only is the walk random, but the graph itself is also random.

An r-regular graph $\mathcal{G}$ is *nice* if [10]

(P'1) $\mathcal{G}$ is connected.

(P'2) The second largest absolute eigenvalue $\lambda^{(2)}_{max}(A_{\mathcal{G}})$ of the adjacency matrix $A_{\mathcal{G}}$ of $\mathcal{G}$ is at most $2\sqrt{r-1}+\epsilon$, where $\epsilon > 0$ is arbitrarily small ($\epsilon = \frac{1}{10}$ is small enough).

(P'3) There are at most $r^{2\sigma}$ vertices on small cycles, where $\sigma = \lfloor \log\log\log n \rfloor$.

(P'4) No pair of small cycles are within distance 3σ of each other.

A sequence of events $\{\mathcal{A}_n\}_{n=0}^{\infty}$ occurs *with high probability* (w.h.p.) in $\mathfrak{G}_n^{(r)}$ when $\lim_{n\to\infty} \mathbb{P}_{\mathfrak{G}_n^{(r)}}(\mathcal{A}_n) = 1$. Theorem 6 in [10] proves that for $r \geq 3$ *constant*, a random r-regular graph is nice w.h.p.. Note that the w.h.p. statement for each property (P'1) to (P'4) means that each property occurs with probability $1 - o(1)$.

Let $\mathcal{G}$ be a realisation of $\mathcal{G}_r^{rnd}$ and let E be an edge chosen u.a.r. from $\mathcal{E}(\mathcal{G})$. We use the results obtained by Cooper and Frieze in [12] to get an asymptotic lower bound for

$\mathbb{P}^{\mu}_{\mathcal{G};\ell}(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E})$. Here μ is *not* necessarily one of the endpoints of a realisation of E but it still denotes the starting point of a random walk of length ℓ on $\mathcal{G}$. In a series of papers [10, 11, 12], Cooper and Frieze developed a technique to estimate the probability that a random walk, starting from $\mu \in \mathcal{V}$, first visits a vertex ν in $\mathcal{G}$ at step $\ell \geq T$, where T is the mixing time. To consider the probability that the random walk first visits an edge e in $\mathcal{G}$, they use a technique of subdividing e by an artificial vertex v_e to form a new graph $\mathcal{H}$. The probability of not visiting e is then given by the probability of not visiting v_e in $\mathcal{H}$.

Step 1: Obtain an upper bound for the mixing time

If $\mathcal{G}$ is nice, then a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$ is ergodic (connectivity of $\mathcal{G}$ and $\lambda^{(2)}_{max}(A_{\mathcal{G}}) < r$ implies ergodicity [10]). If $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ is ergodic, then the stationary distribution $\Pi = [\pi_\nu]_{\nu\in\mathcal{V}}$ of the walk is unique and $\pi_\nu = \frac{d(\nu)}{2m} = \frac{1}{n}$, where $\nu \in \mathcal{V}$. Here $d(\nu) = r$ is the degree of ν and $m = |\mathcal{E}| = \frac{nr}{2}$. Let $\mathcal{G}$ be nice and $P^{\mu\nu}_{\mathcal{G}}(\ell)$ be the ℓ-step transition probability of the random walk on $\mathcal{G}$ from μ to ν. The mixing time $T(\epsilon_1)$ is defined by

$$T(\epsilon_1) = \min\left\{\ell : \sup_{\mu\in\mathcal{V}} \frac{1}{2}\sum_{\nu\in\mathcal{V}}\left|P^{\mu\nu}_{\mathcal{G}}(\ell) - \frac{1}{n}\right| \leq \epsilon_1\right\}.$$

Let $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n$ denote the eigenvalues of the transition probability matrix $P_{\mathcal{G}}$ and $\lambda_{max} = \{|\lambda_i| : i \geq 2\}$. Then for $\ell \geq 1$ [27](see Theorem 12.3), if $\mathcal{G}$ is nice,

$$\left|P^{\mu\nu}_{\mathcal{G}}(\ell) - \frac{1}{n}\right| \leq \lambda^{\ell}_{max}. \tag{14}$$

It was also shown in [17](Theorem 1.1) that, for a random r-regular graph on n

vertices, where $r \geq 3$, and for fixed $\epsilon > 0$, there exists a constant c such that

$$\mathbb{P}_{\mathfrak{G}_n^{(r)}}\left(\lambda_{max} \leq \frac{2\sqrt{r-1}+\epsilon}{r}\right) \geq 1 - \frac{c}{n^K},$$

where $K = \left\lceil \frac{\sqrt{r-1}+1}{2} \right\rceil - 1$. Note that for $r \geq 3$, we have $\frac{2\sqrt{r-1}}{r} \leq \frac{2\sqrt{2}}{3} \leq \frac{29}{30}$. Hence, when $r \geq 3$, one can take $\lambda_{max} \leq \frac{29}{30}$ w.h.p. with respect to n (using $\epsilon \leq 0.0715$). Then by Inequality (14), one can take $T(\epsilon_1) = T \geq 120 \log n$ such that for $\ell \geq T$,

$$\max_{\mu,\nu} \left| P_{\mathcal{G}}^{\mu\nu}(\ell) - \frac{1}{n} \right| \leq \left(\frac{29}{30}\right)^{120 \log n} \leq e^{-4 \log n} = O\left(\frac{1}{n^4}\right),$$

and the distribution of the walk is said to be in *near stationarity* [12].

Step 2: Use the estimated probability of missing a nice edge

Cooper and Frieze gave an estimate of the probability that a random walk, starting from $\mu \in \mathcal{V}$, first visits a vertex $\nu \in \mathcal{V}$, at step $\ell \geq T$ (see [11], Lemma 4). Using this estimate, the probability that the walk does not visit ν at steps $T, T+1, \ldots, \ell$ is obtained (see [11], Corollary 5). Then, as mentioned earlier in this section, the technique of subdividing a given edge e by an artificial vertex v_e is used to obtain the probability that the walk does not visit e at step $T, T+1, \ldots, \ell$. This is given in Lemma 12 below. From now on, we use the expression "w.h.p." to mean with probability $1 - o(\log^{c_1} n)$ for any positive constant c_1.

To simplify the proof of Lemma 12, Cooper and Frieze introduced the notion of a *nice vertex*. For a graph $\mathcal{G}$, a vertex is nice [12] if it is at distance at least $d = \epsilon_2 \log_r n$, for some sufficiently small ϵ_2, from any small cycle $\mathcal{C}$. A cycle is small if $|\mathcal{C}| \leq d$.

Furthermore, it was given in [12] that w.h.p. with respect to the probability space of random r-regular graphs,

1. there are at most $n^{2\epsilon_2}$ vertices that are not nice, and

2. there are no two small cycles within distance d of each other.

An edge $e = (\nu_1, \nu_2) \in \mathcal{E}$ is nice if *both* ν_1, ν_2 are nice [12].

Let $\mathcal{U}_T(\ell)$ denote the set of unvisited edges at step $\ell \geq T$, starting from time T. The following lemma gives the probability that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$, where $\mu \in \mathcal{V}$, does not visit a nice edge $e = (\nu_1, \nu_2) \in \mathcal{E}$ during steps $T, \ldots, \ell$.

Lemma 12. ([12](Lemma 13b)) *For $r \geq 3$ and $\ell \geq \log^3 n$, the following expression holds with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability*

$$\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(e \in \mathcal{U}_T(\ell) | e \in \mathcal{E}(\mathcal{G}) \text{ and } e \text{ is nice}\right) = (1 + o(1))\, e^{-\frac{2\ell}{\rho r n}},$$

where $\mu \in \mathcal{V}$ and $\rho = \frac{r-1}{r-2}$. That is, for any positive constant c_1,

$$\mathbb{P}_{\mathcal{G}_n^{(r)}}\left\{\mathcal{G} : \mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(e \in \mathcal{U}_T(\ell) | e \in \mathcal{E}(\mathcal{G}) \text{ and } e \text{ is nice}\right) = (1 + o(1))\, e^{-\frac{2\ell}{\rho r n}}\right\} = 1 - o(\log^{-c_1} n).$$

Here the expression $o(1)$ is with respect to n.

Step 3: Obtain a lower bound for the probability of visiting a nice edge

Theorem 13. *For $r \geq 3$, $\ell \geq \log^3 n$ and for some sufficiently small ϵ_2, the following lower bound holds with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability*

$$\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E}\right) \geq \left[1 - (1 + o(1))\, e^{-\frac{2\ell}{\rho r n}}\right]\left(1 - \frac{2n^{2\epsilon_2}}{n}\right),$$

where $\rho = \frac{r-1}{r-2}$ and the expression $o(1)$ is with respect to n. That is, for any positive constant c_1,

$$\mathbb{P}_{\mathcal{G}_n^{(r)}} \left\{ \mathcal{G} : \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} \right) \geq \left[1 - (1 + o(1))\, e^{-\frac{2\ell}{\rho r n}} \right] \left(1 - \frac{2n^{2\epsilon_2}}{n} \right) \right\} = 1 - o(\log^{-c_1} n).$$

Proof. Let $\mathcal{G}$ be a realisation of $\mathcal{G}_r^{rnd}$. Let $\mathcal{U}(\ell)$ be the set of unvisited edges at step ℓ. Let $\mathbb{1}(\cdot)$ be an indicator function. We denote the set of nice edges in $\mathcal{G}$ by $\mathcal{E}^{nice}(\mathcal{G})$. For $\mu \in \mathcal{V}(\mathcal{G})$, $1 - \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} | E = e \right) = \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(E \in \mathcal{U}(\ell) | E = e \right)$. Note that $\mathcal{U}(\ell) \subset \mathcal{U}_T(\ell)$. Thus, we have

$$\begin{aligned}
&\mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} \right) \\
&= \sum_{e \in \mathcal{E}(\mathcal{G})} \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} | E = e \right) \mathbb{1}(e \text{ is nice}) \left(\frac{nr}{2} \right)^{-1} + \\
&\quad \sum_{e \in \mathcal{E}(\mathcal{G})} \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} | E = e \right) \mathbb{1}(e \text{ is not nice}) \left(\frac{nr}{2} \right)^{-1} \\
&\geq \sum_{e \in \mathcal{E}(\mathcal{G})} \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} | E = e \right) \mathbb{1}(e \text{ is nice}) \left(\frac{nr}{2} \right)^{-1} \\
&= \frac{2}{nr} \sum_{e \in \mathcal{E}^{nice}(\mathcal{G})} \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E} | E = e \right) \\
&= \frac{2}{nr} \sum_{e \in \mathcal{E}^{nice}(\mathcal{G})} \left[1 - \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(E \in \mathcal{U}(\ell) | E = e \right) \right] \\
&\geq \frac{2}{nr} \sum_{e \in \mathcal{E}^{nice}(\mathcal{G})} \left[1 - \mathbb{P}^{\mu}_{\mathcal{G};\ell} \left(E \in \mathcal{U}_T(\ell) | E = e \right) \right].
\end{aligned}$$

We know that for sufficiently small ϵ_1, with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability there are at most $n^{2\epsilon_1}$ vertices that are not nice. Thus, for sufficiently small ϵ_1, with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability there are *at most* $n^{2\epsilon_1} r$ edges that are not nice, that is, $|\mathcal{E}^{nice}(\mathcal{G})| \geq \frac{nr}{2} - n^{2\epsilon_1} r$.

Then by Lemma 12, for $\ell \geq \log^3 n$, we derive that with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability,

$$\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E}\right) \geq \left[1 - (1 + o(1))\, e^{-\frac{2\ell}{\rho r n}}\right]\left(1 - \frac{2n^{2\epsilon_1}}{n}\right),$$

where $\rho = \frac{r-1}{r-2}$ and the expression $o(1)$ is with respect to n. The above lower bound holds with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability because $\mathcal{G}$ is nice with high $\mathbb{P}_{\mathcal{G}_n^{(r)}}$-probability. $\square$

Note that in Theorem 13, the lower bound for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;E}\right)$, where $\mu \in \mathcal{V}$, is a qualitative bound with respect to n.

Section 4, Section 5 and this section deal with the coverage probability of a given edge in a graph $\mathcal{G}$, specifically when $\mathcal{G}$ is a regular graph. We hope that the coverage probability of an edge could serve as a building block of getting the coverage probability of a prime which consists of more than one edge.

8 Prime factorisation of a walk

A necessary condition for a prime in a graph $\mathcal{G}$ to be recoverable from a random walk $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ is that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ needs to visit all edges of the prime. Given that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ visits all edges of a prime, what is the probability that the prime can be recovered from $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ by using prime factorisation of a walk? An answer to this question depends on what operation is used to factorise a walk. Using different operations to factorise a walk induces different sets of all primes of a graph. Consequently, one obtains different representations of the walk in terms of primes.

8.1 Operations to factorise a walk

Here we use the nesting product as the operation associated to the prime factorisation of a walk.

(a) Nesting product: The set of all primes in $\mathcal{G}$ induced by the nesting product, denoted by $\chi(\mathcal{G})$, is the set of all simple paths and simple cycles in $\mathcal{G}$. The definitions of a simple path and a simple cycle have been given in Section 2. On a finite graph $\mathcal{G}$, the set of all simple paths and simple cycles in $\mathcal{G}$ is finite. The prime factorisation of a walk using the nesting product gives directly a simple path between some $\nu_1, \nu_2 \in \mathcal{V}(\mathcal{G})$ that have been visited by the walk. If the walk visits a cycle from a vertex to itself, this cycle could also be obtained from the prime factorisation. As mentioned in Section 1, simple paths and simple cycles in $\mathcal{G}$ can be used to obtain properties of $\mathcal{G}$.

There are alternative operations to factorise a walk. Each operation induces a set of what could be called "primes" of a graph. We only give some examples and we shall not go into detail about each operation introduced below. A proof of properties of each operation can be found in [18].

(b) Concatenation: The set of all "primes" in $\mathcal{G}$ induced by concatenation, denoted by $\chi^{con}(\mathcal{G})$, is the set of all edges $\mathcal{E}(\mathcal{G})$ in $\mathcal{G}$. $\mathcal{E}(\mathcal{G})$ is a finite set when $\mathcal{G}$ is finite. Comparing to other operations, concatenation gives the smallest set of all "primes" in $\mathcal{G}$. However, concatenation does not facilitate obtaining properties of $\mathcal{G}$ through a walk on $\mathcal{G}$ in the following sense. The prime factorisation of a walk using concatenation gives the set of edges which are visited by the walk. Each edge obtained by the factorisation is like a piece of a puzzle. Assume that the walk visits two vertices ν_1 and ν_2 in $\mathcal{G}$. In order to know a simple path between ν_1 and ν_2 that the walk has covered, one might have to rearrange the edges obtained by the factorisation, just like putting pieces of a part of a puzzle together to see what the part is.

(c) Self-concatenation: A primitive orbit is a cycle C_0 which is not a multiple of a shorter cycle, that is, C_0 is primitive if and only if there is no other cycle C such that $C_0 = C^k$. The set of all "primes" in $\mathcal{G}$ induced by self-concatenation, denoted by $\chi^{scon}(\mathcal{G})$, is the set of all primitive orbits in $\mathcal{G}$. Self-concatenation can only be used to factorise a walk which is a cycle. Moreover, if at least two cycles in $\mathcal{G}$ have at least one vertex in common, $\chi^{scon}(\mathcal{G})$ is a countably infinite set. A countably infinite set of all "primes" of a finite graph $\mathcal{G}$ does not give a good summary of $\mathcal{G}$.

From now on, we say the prime factorisation of a walk to mean the prime factorisation of a walk using the nesting product.

8.2 Nesting product and prime factorisation of a walk

The prime factorisation of a given walk, developed recently by Giscard *et al.* in [19], is akin to the factorisation of an integer into prime numbers. In [19], the authors showed that all walks on arbitrary finite weighted (multi) digraphs obey several non-trivial properties that are largely independent of the graphs on which the walks take place. Foremost amongst these properties is the existence and uniqueness of the factorisation of a given walk into a product of primes. As mentioned above, the set of all primes in a graph $\mathcal{G}$ induced by the nesting product is the set of all simple paths $\mathcal{P}_{\mathcal{G}}$ and the set of all simple cycles $\mathcal{C}_G$ in $\mathcal{G}$.

The nesting product, denoted by $\odot$, is a novel product operation between walks.

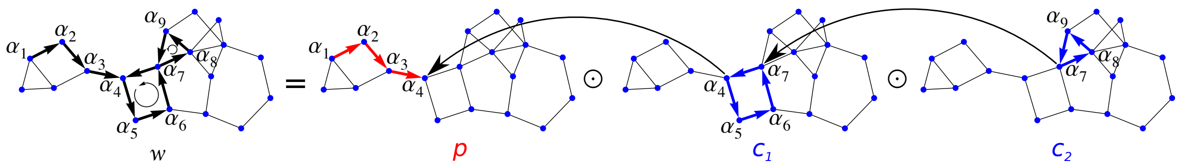


Figure 5: An example of the operation of nesting. The walk w (black arrows) can be obtained by inserting the simple cycle c_2 (blue arrows) into the simple cycle c_1 (blue arrows) and then into the simple path p (red arrows), that is $w = p \odot (c_1 \odot c_2)$. This figure is taken from [19].

It essentially consists of inserting walks into one another and follows the so-called canonical property defined below.

Definition 3. (Canonical property, [19](Definition 1)) *Let* $w_1 = (\mu_0\mu_1 \dots \nu \dots \mu_p)$ *and* $w_2 = (\nu\nu_1\nu_2 \dots \nu_q\nu)$*, where* $w_1, w_2 \in \mathfrak{W}_\mathcal{G}$ *and* $\mu_i = \nu$ *is the last appearance of* ν *in* w_1*. Then the pair* (w_1, w_2) *is canonical if and only if one of the following conditions holds:*

1. *w_1 and w_2 are cycles from ν to itself; or*

2. *no vertex in w_2 other than ν is visited by w_1 before the last appearance of ν in w_1.*

Definition 4. (Nesting product, [19](Definition 2)) *Let* $w_1 = (\mu_0\mu_1 \dots \nu \dots \mu_p)$ *and* $w_2 = (\nu\nu_1\nu_2 \dots \nu_q\nu)$*, where* $w_1, w_2 \in \mathfrak{W}_\mathcal{G}$*, and assume that the pair* (w_1, w_2) *is canonical. The operation of nesting* $\odot : \mathfrak{W}_\mathcal{G} \cup \{0\} \times \mathfrak{W}_\mathcal{G} \cup \{0\} \to \mathfrak{W}_\mathcal{G} \cup \{0\}$ *is defined by*

$$(w_1, w_2) \mapsto w_1 \odot w_2 = (\mu_0\mu_1 \dots \nu\nu_1\nu_2 \dots \nu_q\nu \dots \mu_p).$$

As a simple example illustrating nesting, consider the walk

$$w = (\alpha_1\alpha_2\alpha_3\alpha_4\alpha_5\alpha_6\alpha_7\alpha_8\alpha_9\alpha_7\alpha_4)$$

represented in Fig. 5. This walk is obtained by inserting the simple cycle $c_2 = (\alpha_7\alpha_8\alpha_9\alpha_7)$ into the simple cycle $c_1 = (\alpha_4\alpha_5\alpha_6\alpha_7\alpha_4)$ and then into the simple path $p = (\alpha_1\alpha_2\alpha_3\alpha_4)$, i.e.

$$w = p \odot (c_1 \odot c_2) = (\underbrace{\alpha_1\alpha_2\alpha_3\alpha_4}_{p}\,\underbrace{\alpha_5\alpha_6\,\overbrace{\alpha_7\alpha_8\alpha_9\alpha_7}^{c_2}\,\alpha_4}_{c_1\odot c_2}).$$

Due to the canonical property, it was shown in [19] that the nesting operation, which is non-commutative and non-associative, induces a unique representation for any walk as a product of primes. It then leads to important notions of divisibility of walks, irreducible walks and primes [19].

Definition 5. ([19](Definition 4)) *A walk $w' \in \mathfrak{W}_\mathcal{G}$ divides a walk $w \in \mathfrak{W}_\mathcal{G}$, written as $w'|w$, if and only if there exist $w'_1, w'_2 \in \mathfrak{W}_G$ such that $w = (w'_1 \odot w') \odot w'_2$ or $w = w'_1 \odot (w' \odot w'_2)$.*

Definition 6. ([19]) *A walk $w \in \mathfrak{W}_\mathcal{G}$ is irreducible if there exists $w' \in \mathfrak{W}_\mathcal{G}$ such that $w'|w$, then either $w' = (\mu)$ (a trivial walk), where $\mu \in \mathcal{V}$, or $w = w'$.*

Definition 7. ([19]) *A walk $w \in \mathfrak{W}_\mathcal{G}$ is a prime if and only if for any $w'_1, w'_2 \in \mathfrak{W}_\mathcal{G}$ such that $w|w'_1 \odot w'_2$ then $w|w'_1$ or $w|w'_2$.*

Hence edges are primes by Definition 7. The authors [19] showed that an irreducible walk on $\mathcal{G}$ is a prime and the set of all primes in $\mathcal{G}$ is exactly $\mathcal{P}_\mathcal{G} \cup \mathcal{C}_\mathcal{G}$. Based on the canonical property, for each walk w, there exists a unique factorisation of w into a nesting product of some primes and we call these primes the prime factors of w.

Theorem 14. ([19](Proposition 1)) *A walk $w \in \mathfrak{W}_\mathcal{G}$ is irreducible if and only if w is a prime in $\mathcal{G}$. The set of all primes in $\mathcal{G}$ is exactly $\mathcal{P}_\mathcal{G} \cup \mathcal{C}_\mathcal{G}$.*

Theorem 15. (Unique prime factorisation, [19](Theorem 1)) *Any walk $w \in \mathfrak{W}_{\mathcal{G}}$ factorises uniquely into nesting products of some primes. Consequently, the set of prime factors $S(w)$ of a walk w is uniquely determined by the walk.*

An algorithm producing the prime factorisation of a walk was given in [19]. As an example of the prime factorisation method, consider the walk on four vertices $w = (133112343442333)$. The algorithm of [19] factorises this walk as follows.

1. Traverse the walk from the first vertex to the last appearance of the first vertex (in bold font), **13311**2343442333, and write the walk as, $w = (133112343442333) = (12343442333) \odot (\mathbf{13311})$.

2. Regard the first term in the nesting product as a new walk and proceed as in Step 1, $(1\mathbf{2343442}333) \odot (13311) = (12333 \odot \mathbf{2343442}) \odot (13311)$.

3. For a closed walk, i.e. 13311, traverse the walk from the first vertex to the penultimate appearance of the first vertex, **1331**1, and write the walk as $\mathbf{1331}1 = \mathbf{1331} \odot 11$. Thus, $(12333 \odot 2343442) \odot (13311) = (12333 \odot 2343442) \odot (\mathbf{1331} \odot 11)$.

4. Further factorise the walk w as in Step 2 and Step 3 until each term in the nesting product is either a simple path or a simple cycle, $(12333 \odot 2343442) \odot (1331 \odot 11) = \big((123 \odot \mathbf{33^2}) \odot (23442 \odot \mathbf{343})\big) \odot \big((131 \odot \mathbf{33}) \odot 11\big) = \Big((123 \odot 33^2) \odot \big((2342 \odot \mathbf{44}) \odot 343\big)\Big) \odot \big((131 \odot 33) \odot 11\big)$. Here, $\alpha\alpha^n$ represents the nesting of the cycle $\alpha\alpha$ with itself n times.

5. Return the unique prime factorisation

$$w = \Big((123 \odot 33^2) \odot \big((2342 \odot 44) \odot 343\big)\Big) \odot \big((131 \odot 33) \odot 11\big).$$

Let w denote a walk, Q denote a candidate list of walks and $|\mu_1|(w)$ denote the number of times that vertex μ_1 appears in w. Now, we provide the prime factorisation algorithm, which is a simplified version of [19]; while the prime factorisation algorithm given in [19] cares about the order of prime factors in the representation, we only need the set of prime factors.

Algorithm 1 Prime factorisation of w

$Q \leftarrow w$

$R \leftarrow \emptyset$

$j \leftarrow 1$

while $Q \neq \emptyset$ **do**

  Choose the first entry $w_1 = (\mu_1 \mu_2 \dots \mu_{\ell'})$ in Q and remove it from Q

  if $\mu_1 = \mu_{\ell'}$ **then**

    if $|\mu_1|(w_1) \geq 3$ **then**

      Append $c_1, \dots, c_k$ to Q, where each cycle c_i, $i = 1, \dots, k$, is a walk from the ith appearance of μ_1 until the $(i+1)$th appearance of μ_1 in w_1

    else if $|\mu_1|(w_1) = 2$ and $i' = \min\{i : 2 \leq i \leq \ell' - 2, |\mu_i|(w_1) \geq 2\}$ **then**

      Append $w_1^{(1)} = \left(\mu_1 \dots \mu_{i'_{start}}\right)$ to Q, where $\mu_{i'_{start}}$ is the first appearance of $\mu_{i'}$

      Append $w_1^{(2)} = \left(\mu_{i'_{end}} \dots \mu_1\right)$ to Q, where $\mu_{i'_{end}}$ is the last appearance of $\mu_{i'}$

      Append $c = \left(\mu_{i'_{start}} \dots \mu_{i'_{end}}\right)$ to Q

    else

      Append w_1 to R

    end if

  else if $\mu_1 \neq \mu_{\ell'}$ **then**

if $i' = \min\{i : 1 \leq i \leq \ell' - 1, |\mu_i|(w_1) \geq 2\}$ **then**

Append $w_1^{(1)} = \left(\mu_1 \ldots \mu_{i'_{start}}\right)$ to Q, where $\mu_{i'_{start}}$ is the first appearance of $\mu_{i'}$

Append $w_1^{(2)} = \left(\mu_{i'_{end}} \ldots \mu_{\ell'}\right)$ to Q, where $\mu_{i'_{end}}$ is the last appearance of $\mu_{i'}$

Append $c = \left(\mu_{i'_{start}} \ldots \mu_{i'_{end}}\right)$ to Q

else

Append w_1 to R

end if

end if

end while

return R

8.3 Some observations on the method of prime factorisation

Firstly, we give a necessary and sufficient condition for the edges of a prime to be contained in the prime factors of a walk.

Lemma 16. *Let w be a walk on a graph $\mathcal{G}$, and let $S(w) = \{\chi_1, \chi_2, \ldots, \chi_n\}$ be the set of prime factors of w, for some $n \geq 1$. Let $\chi_{\mathcal{G}}^{\alpha\alpha'}$ be a prime on $\mathcal{G}$ from α to α', where $\alpha, \alpha' \in \mathcal{V}$. We have $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \subseteq \mathcal{E}(w)$ if and only if $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \cap (\cup_{i=1}^{n} \mathcal{E}(\chi_i)) = \mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'})$.*

Proof. If w visits all the edges of $\chi_{\mathcal{G}}^{\alpha\alpha'}$, i.e. $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \subseteq \mathcal{E}(w)$, there exists some subset $\{\chi_{j_1}, \chi_{j_2}, \ldots, \chi_{j_m}\}$ of $S(w)$, where $1 \leq j_1 < j_2 < \ldots < j_m \leq n$, such that the edges in $\chi_{\mathcal{G}}^{\alpha\alpha'}$ are contained in the prime factors in this subset, i.e. $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) = \mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \cap \left(\cup_{i'=1}^{m} \mathcal{E}\left(\chi_{j_{i'}}\right)\right)$. Consequently, we have $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \cap (\cup_{i=1}^{n} \mathcal{E}(\chi_i)) = \mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'})$.

The condition $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \subseteq \mathcal{E}(w)$ is violated as soon as the walk w does not visit at least one edge contained in $\chi_{\mathcal{G}}^{\alpha\alpha'}$. Thus there exists at least one edge $e \in \mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'})$ such that $e \notin \cup_{i=1}^{n} \mathcal{E}(\chi_i)$. Consequently, $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \cap (\cup_{i=1}^{n} \mathcal{E}(\chi_i)) \neq \mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'})$. $\square$

If $\chi_{\mathcal{G}}^{\alpha\alpha'} = (\alpha, \alpha') \in \mathcal{E}$ and w visits (α, α'), then (α, α') appears in some prime factor of w. It is worth pointing out that Lemma 16 remains true regardless of the order in which w visits the edges of a prime. However we remark that when $\chi_{\mathcal{G}}^{\alpha\alpha'}$ consists of more than one edge, i.e. $\chi_{\mathcal{G}}^{\alpha\alpha'} = (\alpha\alpha_1 \ldots \alpha_{\ell'-1}\alpha')$ and $\ell' > 1$, and when the condition $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \subseteq \mathcal{E}(w)$ is true, one can directly read the prime $\chi_{\mathcal{G}}^{\alpha\alpha'}$ off the prime factorisation of w only if there exists some i such that $\mathcal{E}(\chi_{\mathcal{G}}^{\alpha\alpha'}) \subseteq \mathcal{E}(\chi_i)$ and $\chi_i = (\beta \ldots \alpha\alpha_1 \ldots \alpha_{\ell'-1}\alpha' \ldots \beta')$ or $\chi_i = (\beta \ldots \alpha'\alpha_{\ell'-1} \ldots \alpha_1\alpha \ldots \beta')$ ($\beta = \beta'$ are not necessarily distinct), where $\chi_i \in S(w)$. In this case, we say $\chi_{\mathcal{G}}^{\alpha\alpha'}$ is contained in a prime factor χ_i of w. To illustrate this, we define what a partially canonical walk is and give three conditions for some prime factors of a walk to contain $\chi_{\mathcal{G}}^{\alpha\alpha'}$.

Definition 8. (Partially canonical walk) *Let $\chi_{\mathcal{G}}^{\alpha\alpha'} = (\alpha\alpha_1 \ldots \alpha_{\ell'-1}\alpha')$, where $\ell' > 1$ and $\alpha, \alpha' \in \mathcal{V}$ with $\alpha \equiv \alpha_0$, $\alpha' \equiv \alpha_{\ell'}$. Assume that a walk w on $\mathcal{G}$ visits all edges in $\chi_{\mathcal{G}}^{\alpha\alpha'}$ and let $\{\mu_1, \ldots, \mu_k\}$ be the set of distinct vertices visited by w before the first occurrence of α in w. We say that the walk w is a partially canonical walk for $\chi_{\mathcal{G}}^{\alpha\alpha'}$ if the following conditions hold:*

1. *For $1 \leq i \leq \ell'$, the first occurrence of α_i in w happens after the first occurrences of $\alpha, \alpha_1, \ldots, \alpha_{i-1}$ in w.*

2. *If a cycle from $\mu' \in \{\mu_1, \ldots, \mu_k\}$ to itself visits some vertices in $\chi_{\mathcal{G}}^{\alpha\alpha'}$, then the vertices in $\chi_{\mathcal{G}}^{\alpha\alpha'}$ visited by the cycle must appear again in their corresponding order in $\chi_{\mathcal{G}}^{\alpha\alpha'}$ after the cycle in w.*

3. *If a cycle from α_i to itself, where $1 \leq i \leq \ell'$, does not visit any vertex in $\{\alpha, \alpha_1, \ldots, \alpha_{i-1}\}$ but visits some vertices in $\{\alpha_{i+1}, \ldots, \alpha'\}$, then the vertices in $\{\alpha_{i+1}, \ldots, \alpha'\}$ visited by the cycle must appear again in their corresponding order in $\chi_{\mathcal{G}}^{\alpha\alpha'}$ after the cycle in w.*

We also say that the walk w is a partially canonical walk for $\chi_{\mathcal{G}}^{\alpha\alpha'}$ if the three conditions above hold for the reverse sequence $(\alpha'\alpha_{\ell'-1}\ldots\alpha_1\alpha)$.

One can see that the three conditions in Definition 8 for a given prime in a graph $\mathcal{G}$ being recovered from a walk on $\mathcal{G}$ are restrictive.

On the other hand, to get a larger set of primes from a given walk w, we factorise w at each step when we traverse through w. We call this method "*accumulative prime factorisation*" of a walk. We revisit the example given above where $w = (133112343442333)$. Firstly, we traverse through w from the left and obtain a set of "subwalks" of w, which is

$$\{(13), (133), (1331), \ldots, (13311234344233), (133112343442333)\}.$$

We then factorise the subwalk obtained at each step using the prime factorisation method given in [19]. Applying the prime factorisation method in [19] to each subwalk is a routine and we shall not do it here; the example shown in Section 8.2 provides a step-by-step illustration.

We recall that the set of prime factors of w obtained by using the prime factorisation method given in [19] is $\{(11), (33), (44), (123), (131), (343), (2342)\}$. By using the accumulative prime factorisation method, we get an additional set of prime factors $\{(12), (13), (1234)\}$. We recall that the set of prime factors of a walk w' is denoted by $S(w')$. The accumulative prime factorisation of a walk is not monotonic in the sense that the set of prime factors of the walk obtained at step i is not necessarily a subset of the set of prime factors of the walk obtained at step $i + 1$. To see the non-monotonicity of the accumulative prime factorisation method, note, for example, that $S((133)) = \{(13), (33)\} \not\subset S((1331)) = \{(131), (33)\}$, and $S((13311234344)) = \{(11), (33), (44), (131), (343), (1234)\}$, which is not a subset of

$S((133112343442)) = \{(11), (12), (33), (44), (131), (343), (2342)\}$. Although one gets a larger set of primes by using the accumulative prime factorisation method, a probabilistic analysis of the recoverability of a given prime on this method is harder due to non-monotonicity of the method.

9 Exploratory simulations and open problems

For exploratory purposes, we show some simulation results for the following two questions:

1. For a fixed r-regular graph $\mathcal{G}_r$, given that $e \in \mathcal{E}(\mathcal{G}_r)$ and that $\mu \in \mathcal{V}(\mathcal{G}_r)$ is one of the endpoints of e, how does the probability that $\mathcal{W}^{\mu}_{\mathcal{G}_r;\ell}$ visits e behave *with respect to the degree of the vertices* r? Then consider selecting a graph $\mathcal{G}^{rnd}_r$ u.a.r. from the set of r-regular graphs, and for the realisation of $\mathcal{G}^{rnd}_r$, we select a starting point μ of the random walk u.a.r. and an edge is selected u.a.r. from those incident to μ. In this case, how does the coverage probability of an edge behave with respect to the degree of the vertices r?

2. Assume that χ is a prime in a fixed r-regular graph $\mathcal{G}_r$. Let $\ell(\chi)$ denote the length of χ and let $\mu \in \mathcal{V}(\mathcal{G}_r)$ be one of the endpoints of χ. How does the probability that χ appears in the *accumulative* prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G}_r;\ell}$ behave when $\ell(\chi)$ *increases*? Again, consider selecting a graph $\mathcal{G}^{rnd}_r$ u.a.r. from the set of r-regular graphs, and for the realisation of $\mathcal{G}^{rnd}_r$, we select a starting point μ of the random walk u.a.r. and a prime of length ℓ_p is selected u.a.r from the set of primes of length ℓ_p in which μ is one of the endpoints. In this case, how does the probability that a prime of length ℓ_p appears in the accumulative prime factorisation of a realisation of the random walk behave when ℓ_p increases?

Let $\mathcal{B}$ denote the event that we are interested in, such as $\mathcal{B} = \mathcal{A}^{\mu}_{\mathcal{G};\ell;e}$. Let p denote the probability of the event $\mathcal{B}$. We want to estimate p through simulations. Let m be the number of runs used to estimate p. We estimate p by $\hat{p}_m = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{B}_i}$, where $\mathbb{1}_{\mathcal{B}_i}$ is 1 if the event $\mathcal{B}$ occurs in the i-th sample and 0 otherwise. Based on the context in Section 9.1 or Section 9.2, $\mathcal{B}$ varies. Notice that for $1 \le i \le m$, $\mathbb{1}_{\mathcal{B}_i}$ are independent and identically distributed random variables. Hence, we have

$$\mathbb{E}[\hat{p}] = \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{B}_i}\right] = p, \quad \text{and}$$

$$\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{B}_i}\right) = \frac{p(1-p)}{m} \le \frac{1}{4m}.$$

By the Strong Law of Large Numbers, we know that $\hat{p}$ converges to p *almost surely* when $m \to \infty$, that is,

$$\mathbb{P}\left(\lim_{m\to\infty} \hat{p} = p\right) = 1.$$

By the Central Limit Theorem, we also have

$$\frac{\sqrt{m}(\hat{p}-p)}{\sqrt{p(1-p)}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $\xrightarrow{d}$ denotes convergence in distribution and $\mathcal{N}(0,1)$ denotes the standard normal distribution.

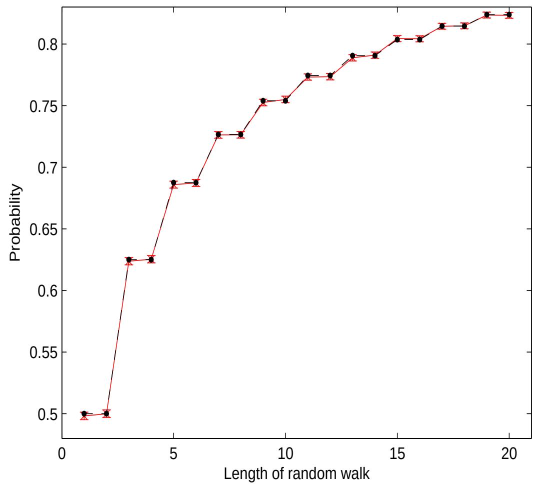
In Section 9.1 and Section 9.2, we use $m = 100,000$ to get reasonable convergence of $\hat{p}_m$. The standard deviation is then estimated by $\sqrt{\frac{\hat{p}(1-\hat{p})}{m}}$. As a sanity check, 95% confidence intervals for p (using normal approximations) are used in the following sections to show the reliability of the estimates of p. All the simulation results shown in the following sections are obtained using MATLAB (version R2012b).

9.1 Coverage probability of an edge with respect to the degree of the vertices

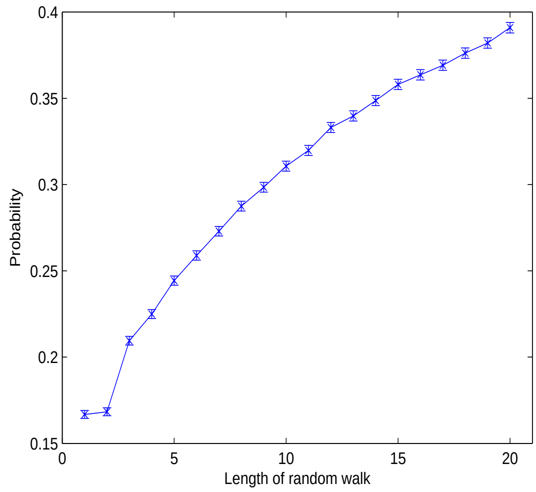
Consider a fixed r-regular graph $\mathcal{G}$. We address $\mathbb{P}^{\mu}_{\mathcal{G};\ell;e}$, in terms of r. First we use a nested scheme by considering the k-nearest neighbour graphs $\mathcal{G}(n,k)$. We say that a graph $\mathcal{G}$ is *nested* in another graph $\mathcal{G}'$ if $\mathcal{V}(\mathcal{G}) = \mathcal{V}(\mathcal{G}')$ and $\mathcal{E}(\mathcal{G}) \subseteq \mathcal{E}(\mathcal{G}')$. Thus if $1 \le k \le k' \le \lfloor \frac{n}{2} \rfloor$, $\mathcal{G}(n,k)$ is nested in $\mathcal{G}(n,k')$. We recall that $\mathcal{G}(n,1)$ is a circle graph and $\mathcal{G}(n, \lfloor \frac{n}{2} \rfloor)$ is the complete graph.

We fix n to be 1,000. For $\mathcal{G} = \mathcal{G}(n,k)$, we select a given edge $e \in \mathcal{E}(\mathcal{G})$. Due to symmetry of $\mathcal{G}$, it does not matter which edge e we consider. The starting point μ of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ is one of the endpoints of e. By symmetry of $\mathcal{G}$, $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ does not depend on which endpoint of e is selected as μ. For each ℓ, we estimate $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ by performing $m = 100,000$ independent $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$ and using $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{A}_i}$, where $\mathbb{1}_{\mathcal{A}_i}$ returns 1 if the i-th random walk on $\mathcal{G}$ visits e, and 0 otherwise. Let us call this simulation procedure "Simulation A". Note that for different values of ℓ, we use a new set of m independent $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$.

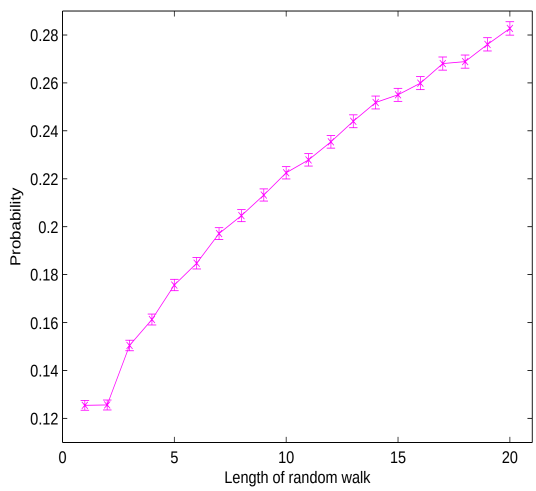
We first check if the simulation scheme produces reliable results by computing 95% confidence intervals. For each graph $\mathcal{G}$ and ℓ, an estimate of a 95% confidence interval for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ is constructed as $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right) \pm 1.96\sqrt{\frac{\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)\left[1-\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)\right]}{m}}$, where $m = 100,000$. For a sanity check, we select four different graphs and estimate 95% confidence intervals for $\ell = 1, 2, \ldots, 20$ for each graph. The estimated 95% confidence intervals for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ are plotted in Fig. 6. We can see that the estimated 95% confidence intervals for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ are narrow. When $\mathcal{G} = \mathcal{C}_{1000}$ or $\mathcal{G} = \mathcal{K}_{1000}$, for



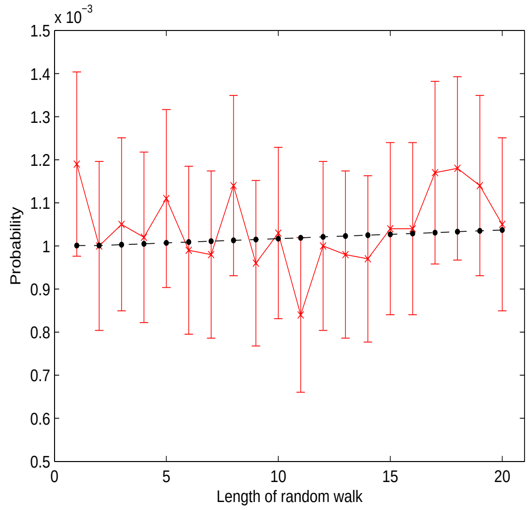
(a) Circle graph C_{1000}



(b) 3-nearest neighbour graph $\mathcal{G}(1000, 3)$



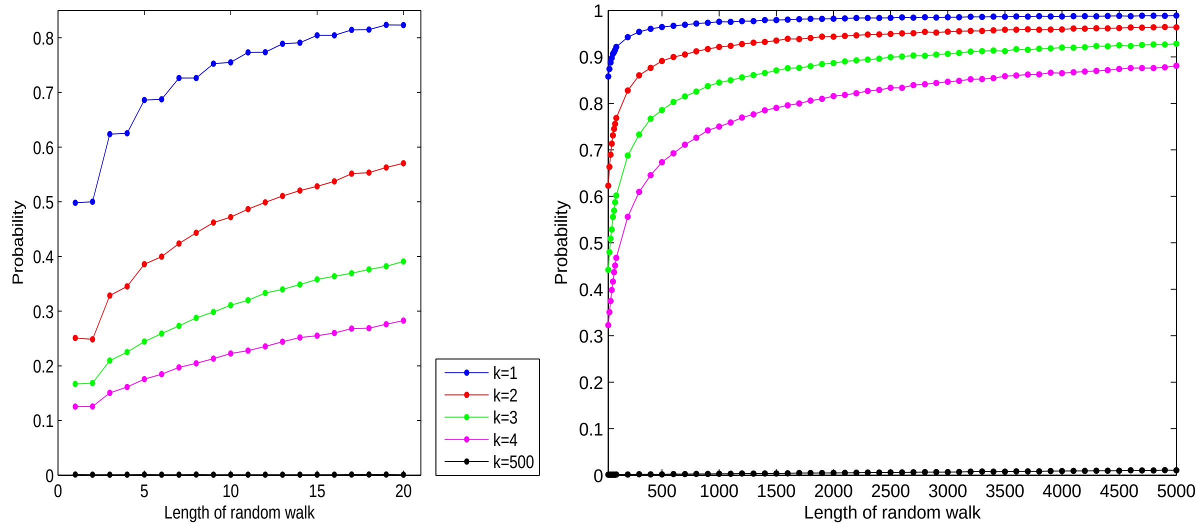
(c) 4-nearest neighbour graph $\mathcal{G}(1000, 4)$



(d) Complete graph $\mathcal{K}_{1000}$

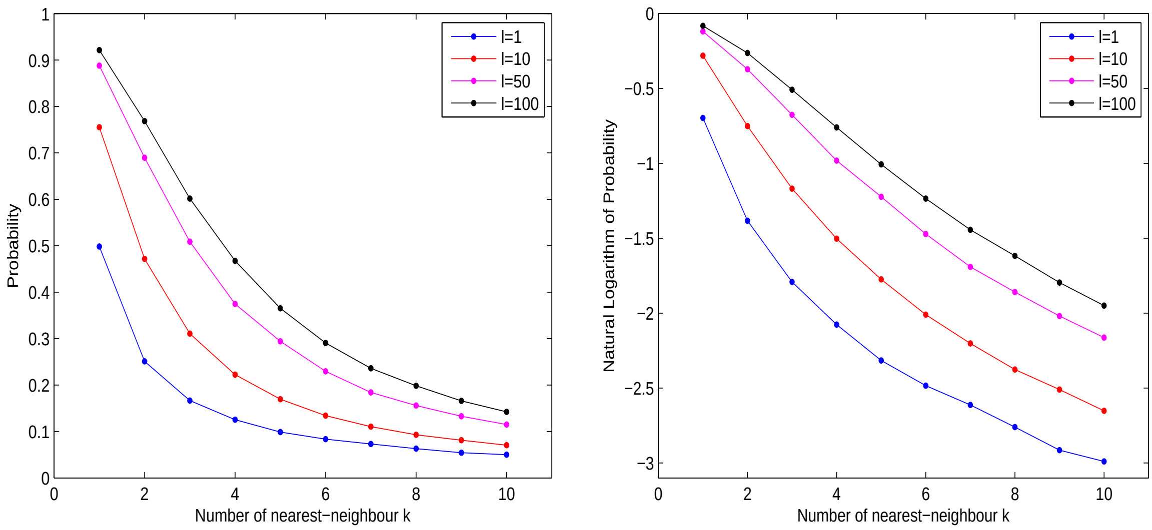
Figure 6: 95% confidence intervals for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ for $\ell = 1, 2, \ldots, 20$. For each graph and for each ℓ, the cross represents $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$. The black dots in (a) and (d) represent the true coverage probabilities. Note that in (d) the probabilities are on a 10^{-3} scale.

each $\ell = 1, 2, \ldots, 20$, we can calculate $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ (using the formula in Theorem 3 or Theorem 4 respectively). For $\mathcal{C}_{1000}$ and $\mathcal{K}_{1000}$, the true probabilities are shown as black dots in Fig. 6a and Fig. 6d respectively. Note that for $\mathcal{G} = \mathcal{C}_{1000}$ or $\mathcal{G} = \mathcal{K}_{1000}$, the true probability $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ lies in the estimated 95% confidence interval for each $\ell = 1, 2, \ldots, 20$.



(a) $\ell \in \{1, 2, \ldots, 20\}$ (b) $\ell \in \{30, 40, \ldots, 100\} \cup \{200, 300, \ldots, 5000\}$

Figure 7: $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ for $\mathcal{G} = \mathcal{G}(n, k)$, where $n = 1000$ and $k = 1, 2, 3, 4, 500$.



(a) (b)

Figure 8: For $\mathcal{G} = \mathcal{G}(n, k)$, (a) $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ with respect to k for fixed ℓ; (b) Natural logarithm of $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ with respect to k for fixed ℓ.

We can see that for each $\mathcal{G} = \mathcal{G}(n,k)$ in Fig. 7a, we have that $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};1}\left(\mathcal{A}^{\mu}_{\mathcal{G};1;e}\right)$ is approximately $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};2}\left(\mathcal{A}^{\mu}_{\mathcal{G};2;e}\right)$. In fact, $\mathbb{P}^{\mu}_{\mathcal{G};1}\left(\mathcal{A}^{\mu}_{\mathcal{G};1;e}\right) = \mathbb{P}^{\mu}_{\mathcal{G};2}\left(\mathcal{A}^{\mu}_{\mathcal{G};2;e}\right)$ as $\mathcal{W}^{\mu}_{\mathcal{G};1}$ either visits e or moves one step away from μ. If $\mathcal{W}^{\mu}_{\mathcal{G};1}$ moves one step away from μ, then $\mathcal{W}^{\mu}_{\mathcal{G};2}$ either moves further away from μ or moves back to one of the endpoints of e without visiting e. For $j = 1, 2, \ldots$, the behaviour that $\mathbb{P}^{\mu}_{\mathcal{G};2j-1}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j-1;e}\right) = \mathbb{P}^{\mu}_{\mathcal{G};2j}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j;e}\right)$ is preserved only when $\mathcal{G}$ is a circle graph.

In Fig. 7 and Fig. 8, we see that for fixed ℓ in this particular nested scheme where $\mathcal{G} = \mathcal{G}(n,k)$, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$ decreases when k increases. The rate of decrease seems to be approximately exponential (see Fig. 8b).

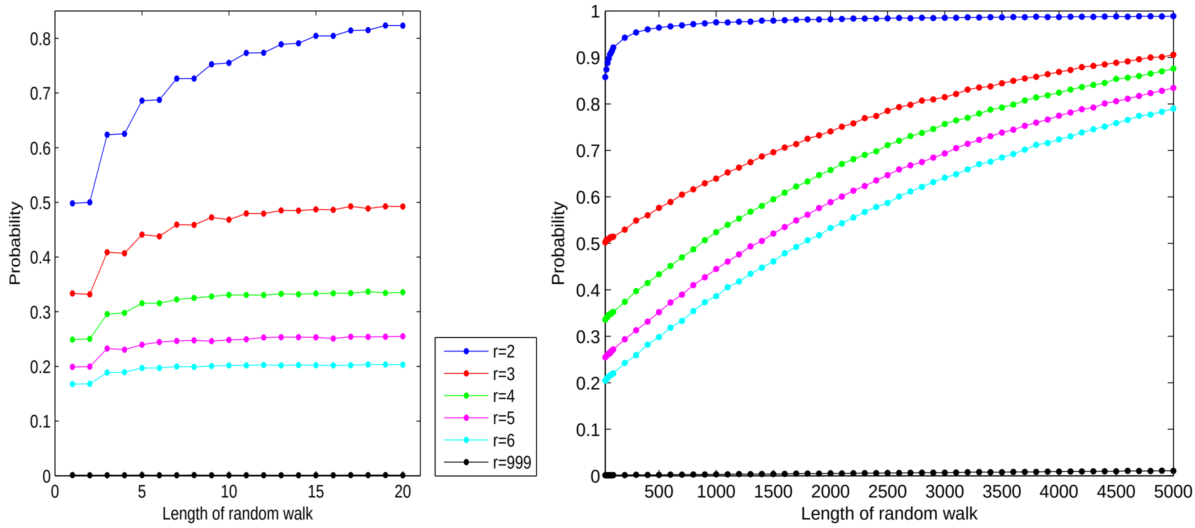
We now consider arbitrary connected r-regular graphs $\mathcal{G}_r$ on n vertices, which are picked u.a.r. from $\mathfrak{G}^{(r)}_n$. Note that the only connected 2-regular graph on n vertices is a circle graph and the only connected $(n-1)$-regular graph on n vertices is the complete graph. Thus for arbitrary connected r-regular graphs, we consider $2 < r < n-1$. Let $n = 1,000$ and $m = 100,000$ (the sample size). For each ℓ and for each sample, we perform the following simulation steps.

1. We select an r-regular graph randomly using the MATLAB function createRandRegGraph and check if the selected graph is connected. If the selected graph is not connected, we re-sample a graph using createRandRegGraph until a connected r-regular graph is obtained.

2. For the sampled graph, we select a vertex μ with probability $\frac{1}{n}$ and select an edge e from those incident to μ uniformly.

3. We perform one random walk of length ℓ starting from μ and check if the walk visits e.

Let us call this simulation procedure "Simulation B".

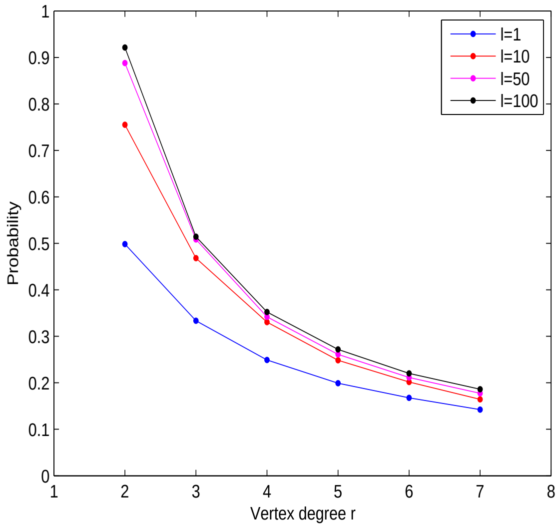


(a) $\ell \in \{1, 2, \ldots, 20\}$ (b) $\ell \in \{30, 40, \ldots, 100\} \cup \{200, 300, \ldots, 5000\}$

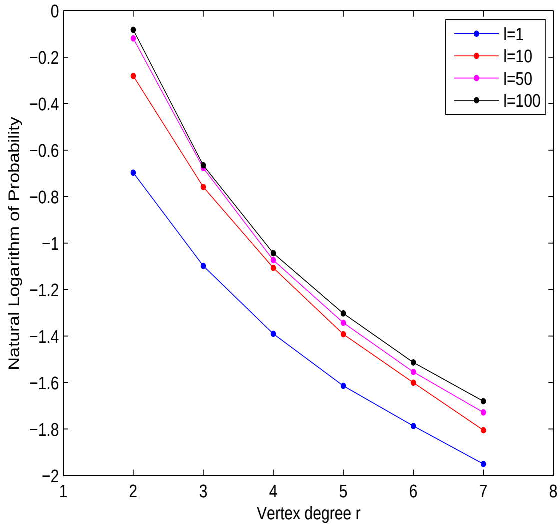
Figure 9: For $n = 1000$, $\hat{p}_1(r, \ell)$ where $r = 3, 4, 5, 6$. The dark blue line and the black line are $\hat{\mathbb{P}}^{\mu}_{\mathcal{C}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;e}\right)$ and $\hat{\mathbb{P}}^{\mu}_{\mathcal{K}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{K}_n;\ell;e}\right)$ respectively.

Repeating this procedure, we sample $m = 100,000$ connected r-regular graphs, then we estimate the probability that on a random connected r-regular graphs, a random walk of length ℓ starting from a randomly chosen vertex μ visits a randomly selected edge incident to μ. For each r and for each ℓ, the estimator that we use is $\hat{p}_1(r, \ell) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{A}_i}$, where $1_{\mathcal{A}_i}$ returns 1 if in the i-th sample, the random walk visits the selected edge, and 0 otherwise. Note that using $m = 100,000$ is not enough to sample "most" connected r-regular graphs on $n = 1,000$ vertices. For $r = 3$, the number of connected 3-regular graphs on 20 vertices is 510,489 [35]. Thus, the simulation results shown here are only for exploratory purposes.

As for the k-nearest neighbour graphs, we can see in Fig. 9 that $\hat{p}_1(r, 1)$ is approximately $\hat{p}_1(r, 2)$. The reason is the same as for k-nearest neighbour graphs. From Fig. 9 and Fig. 10, we see that for fixed ℓ, $\hat{p}_1(r, \ell)$ decreases when r increases. The rate of decrease seems to be approximately exponential (see Fig. 10b).



(a)

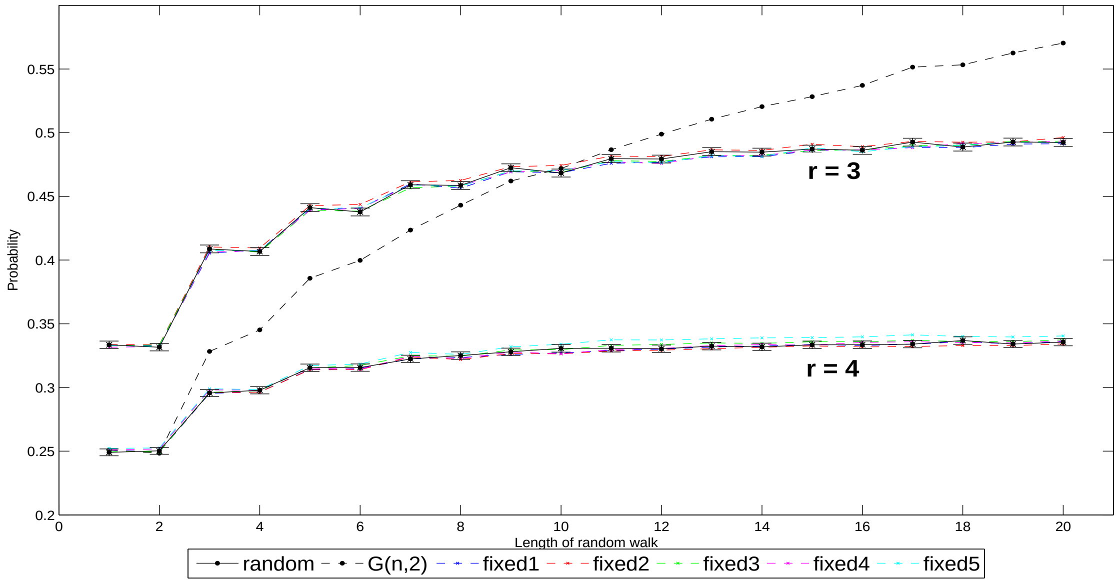


(b)

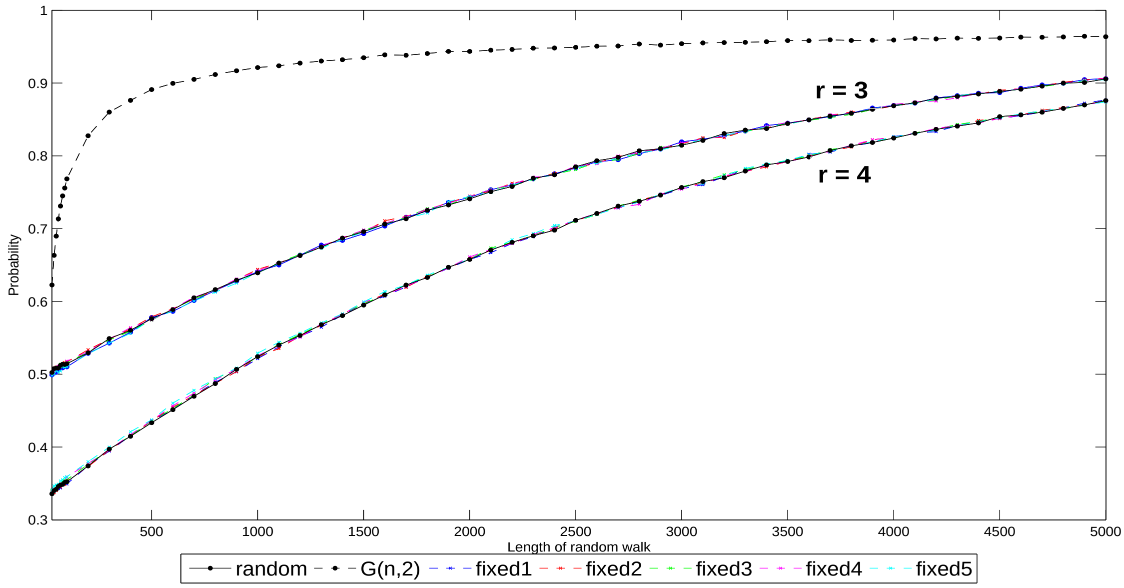
Figure 10: (a) $\hat{p}_1(r, \ell)$ with respect to r for fixed ℓ; (b) Natural logarithm of $\hat{p}_1(r, \ell)$ with respect to r for fixed ℓ.

Notice that $\hat{p}_1(r, \ell)$ tends to grow more slowly with respect to ℓ comparing to k-nearest neighbour graphs. Here is a possible explanation. For $\mathcal{G}(n, k)$, the endpoints of a selected edge lie on some triangle in $\mathcal{G}(n, k)$. A random walk starting from one of the endpoints of the selected edge e is fairly likely to revisit the endpoints of e and then visit e. For $r \geq 3$, $\mathcal{G}_r^{rnd}$ is locally tree-like w.h.p [10]. This means that once a random walk starting from one of the endpoints an edge e moves away from the endpoints of e, typically it takes a fair amount of time for the random walk to revisit one of the endpoints of e and then visit e. To explore this observation further, we sample some connected r-regular graphs from the set of r-regular graphs. For each sampled graph, we perform simulation procedure "Simulation A", as for k-nearest neighbour graphs.

For exploratory purposes, we only consider the case where $r = 3$ or $r = 4$ and $n = 1,000$. Note that there is no k-nearest neighbour graphs with vertex degree 3. We sample 5 connected 3-regular graphs from the set of 3-regular graphs by using the MATLAB function createRandRegGraph. For each $\mathcal{G}_3$ of the sampled regular graphs,



(a) $\ell \in \{1, 2, \ldots, 20\}$



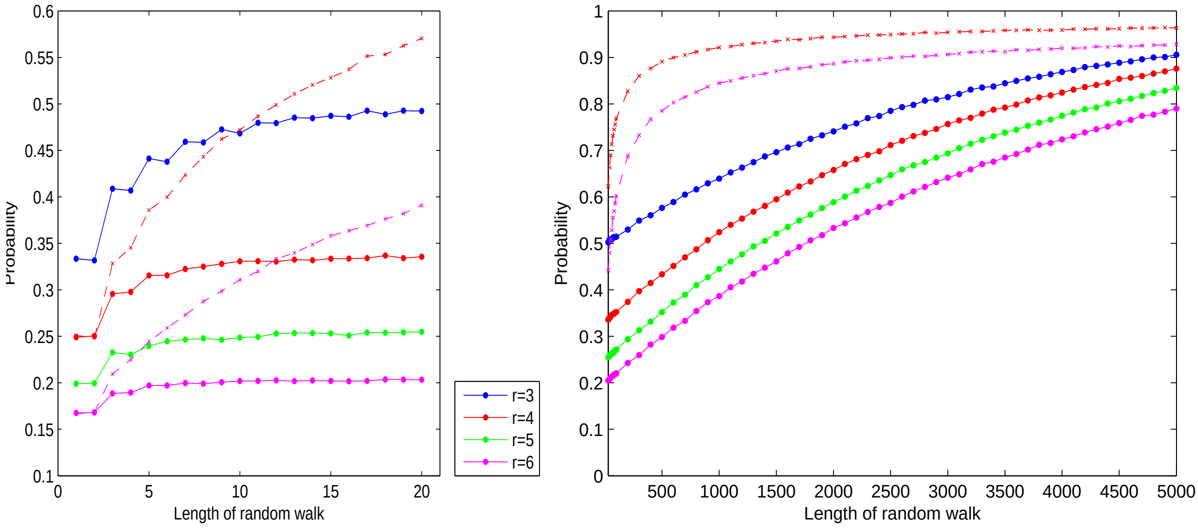
(b) $\ell \in \{30, 40, \ldots, 100\} \cup \{200, 300, \ldots, 5000\}$

Figure 11: Probabilities of covering an edge for some 3-regular graphs and some 4-regular graphs. For $r = 3$ or $r = 4$, the graphs for "fixed1",…,"fixed5" are 5 connected graphs sampled from $\mathfrak{G}_n^{(r)}$. The probabilities for "$\mathcal{G}(n, 2)$", "fixed1",…,"fixed5" are obtained using simulation procedure "Simulation A", i.e. $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;e}\right)$. The probabilities for "random" are obtained using simulation procedure "Simulation B", i.e. $\hat{p_1}(r, \ell)$. For $\hat{p_1}(3, \ell)$ and $\hat{p_1}(4, \ell)$, estimates of 95% confidence intervals for the true coverage probabilities are shown in (a).

we select $\mu \in \mathcal{V}(\mathcal{G}_3)$ with probability $\frac{1}{n}$ and select an edge e from those incident to μ with probability $\frac{1}{r}$. Using simulation procedure "Simulation A", for each ℓ, we run $m = 100,000$ independent $\mathcal{W}^{\mu}_{\mathcal{G}_3;\ell}$ on each $\mathcal{G}_3$ and estimate $\mathbb{P}^{\mu}_{\mathcal{G}_3;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_3;\ell;e}\right)$ by $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_3;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_3;\ell;e}\right) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{A}_i}$, where $\mathbb{1}_{\mathcal{A}_i}$ returns 1 if the i-th random walk on $\mathcal{G}_3$ visits e, and 0 otherwise. For $r = 4$, we also sample 5 connected 4-regular graphs $\mathcal{G}_4$ from the set of 4-regular graphs by using the MATLAB function createRandRegGraph but we also add the case when $\mathcal{G}_4$ is a 2-nearest neighbour graph. Again, for each $\mathcal{G}_4$ and for each ℓ, we perform simulation procedure "Simulation A" to estimate $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_4;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_4;\ell;e}\right)$.

The simulation results are shown in Fig. 11. We can see that for each of the 5 sampled 3-regular graph $\mathcal{G}_3$ and for each ℓ, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_3;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_3;\ell;e}\right)$ is fairly well concentrated. We notice that for each $\mathcal{G}_3$ and for each ℓ, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_3;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_3;\ell;e}\right)$ is close to $\hat{p}_1(3,\ell)$. This could be due to the fact that the sampled graphs have the same property, which are locally tree-like (w.h.p.). Fig. 11 shows the variability between the coverage probabilities for $r = 4$ and shows that the structure of a regular graph *does* affect the coverage probabilities. In particular for $\mathcal{G}(n,2)$, the coverage probabilities tend to grow much faster than the coverage probabilities for the sampled 3-regular graphs and the coverage probabilities for the sampled 4-regular graphs. However, for each of the 5 sampled 4-regular graph $\mathcal{G}_4$ and for each ℓ, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_4;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_4;\ell;e}\right)$ is fairly well concentrated and $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}_4;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}_4;\ell;e}\right)$ is close to $\hat{p}_1(4,\ell)$.

We now compare regular graphs with different vertex degrees. The "dashed-cross" lines in Fig. 12 represent k-nearest neighbour graphs, where $k = 2, 3$, and the solid lines represent random connected r-regular graphs, where $r = 3, 4, 5, 6$. We can see that, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G}(n,k);\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G}(n,k);\ell;e}\right)$ tends to grow faster than $\hat{p}_1(r,\ell)$. Thus for r-regular graphs, we *do not* find a general monotone relationship between the probability of visiting an edge and the degree r.



(a) $\ell \in \{1, 2, \ldots, 20\}$ (b) $\ell \in \{30, 40, \ldots, 100\} \cup \{200, 300, \ldots, 5000\}$

Figure 12: A comparison of the probabilities of covering an edge for different r-regular graphs. The "dashed-cross" lines represent k-nearest neighbour graphs, where $k = 2, 3$ (using simulation procedure "Simulation A"). The solid lines represent random connected r-regular graphs, where $r = 3, 4, 5, 6$ (using simulation procedure "Simulation B").

9.2 Probability of recovering a prime with respect to the length of the prime

We have seen in Section 8.3 that given a walk w of length $\ell \geq 3$, the *accumulative* prime factorisation of w gives more prime factors than using the prime factorisation method developed in [19] to factorise w. Consider $\mathcal{G}$ being a fixed graph. Let χ be a given prime in $\mathcal{G}$ of length $\ell(\chi)$ and let μ be one of the endpoints of χ. How does the probability that χ appears in the accumulative prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ behave when $\ell(\chi)$ increases? We recall that there are different ways to factorise a walk accumulatively. In this section, we only consider factorising a walk w accumulatively from the left end of w (see Section 8.3). By saying that a prime $\chi = (\mu_0\mu_1 \ldots \mu_{\ell(\chi)})$ appears in the accumulative prime factorisation, we mean that χ

appears in some prime factor of a walk obtained by the accumulative prime factorisation in the form of $(\mu_0\mu_1\ldots\mu_{\ell(\chi)})$ or $(\mu_{\ell(\chi)}\ldots\mu_1\mu_0)$. Let us denote the event that χ appears in the accumulative prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ by $\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}$. We shall also call $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ the recoverability probability of χ. Note that for a simple graph $\mathcal{G}$, a walk of length 2 is a prime of length 2 on $\mathcal{G}$. However, a walk of length 3 may not be a prime of length 3 on $\mathcal{G}$. Thus for a regular graph $\mathcal{G}_r$, when $\ell(\chi)=2$, $\mathbb{P}^{\mu}_{\mathcal{G};2}\left(\mathcal{A}^{\mu}_{\mathcal{G};2;\chi}\right)=\frac{1}{r^2}$.

First we consider $\mathcal{G}=\mathcal{G}(n,k)$. We fix n to be 1,000. We select a given prime χ of length 1,2 or 3 in $\mathcal{G}$ deterministically. We shall see later in the simulation results that for some graphs, $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ *depends* on the selection of χ. The starting point of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ is one of the endpoints of χ. Due to symmetry of $\mathcal{G}$, $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ does not depend on which endpoint of χ is selected as μ. For each ℓ, we estimate $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ by performing $m=100,000$ independent $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ on $\mathcal{G}$ and using $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)=\frac{1}{m}\sum_{i=1}^{m}\mathbb{1}_{\mathcal{A}_i}$, where $\mathbb{1}_{\mathcal{A}_i}$ returns 1 if χ appears in some prime factor (using accumulative prime factorisation) of the realisation of the i-th random walk, and 0 otherwise. This is purely exploratory; we only consider $\ell\in\{1,2,\ldots,20\}$ to get a sense of how $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ behaves with respect to $\ell(\chi)$. To show the reliability of the estimate of $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$, we use an estimate of a 95% confidence interval for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ constructed by $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)\pm 1.96\sqrt{\frac{\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)\left[1-\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)\right]}{m}}$.

Fig. 13 shows that for $\mathcal{G}=\mathcal{C}_n$, when $\ell(\chi)=2h-1$ where $h=1,2$ (i.e. $\ell(\chi)=1$ or 3), we have that $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};2j-1}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j-1;\chi}\right)$ is close to $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};2j}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j;\chi}\right)$ for $j\geq 2h-1$. When $\ell(\chi)=2$, we have that $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};2j}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j;\chi}\right)$ is close to $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};2j+1}\left(\mathcal{A}^{\mu}_{\mathcal{G};2j+1;\chi}\right)$ for $j\geq 1$. To understand this behaviour, let us take $\ell(\chi)=2$ as an example. χ either appears in some prime factor of the accumulative prime factorisation of the realisation of $\mathcal{W}^{\mu}_{\mathcal{G};2}$ or not. If χ does not appear in the prime factors of the accumulative prime factorisation

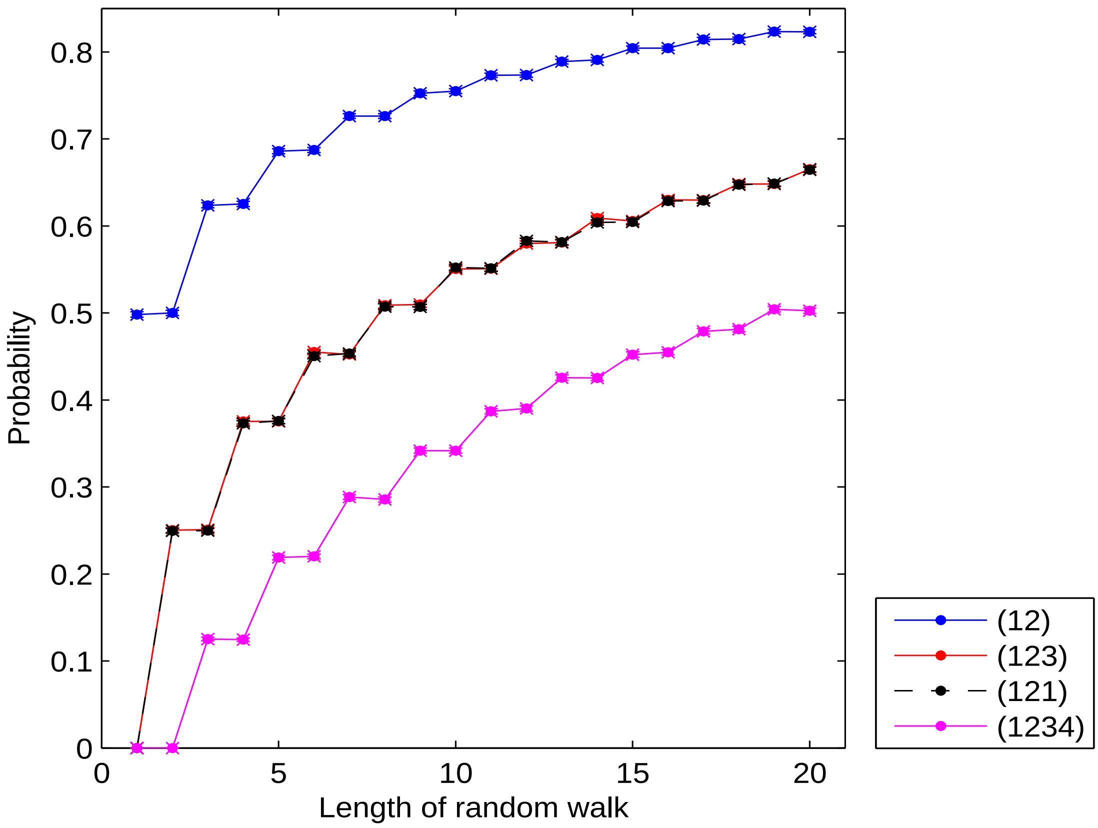


Figure 13: $\hat{\mathbb{P}}^{\mu}_{\mathcal{C}_{1000};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_{1000}\ell;\chi}\right)$ for $\chi \in \{(12),(123),(121),(1234)\}$, where $\ell = 1, 2, \ldots, 20$. For each χ and for each ℓ, an estimate of a 95% confidence interval for $\mathbb{P}^{\mu}_{\mathcal{C}_{1000};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_{1000}\ell;\chi}\right)$ is also shown.

of the realisation of $\mathcal{W}^{\mu}_{\mathcal{G};2}$, then χ does not appear in the prime factors of the prime factorisation of the realisation of $\mathcal{W}^{\mu}_{\mathcal{G};3}$ (one just has to factorise the realisation of $\mathcal{W}^{\mu}_{\mathcal{G};3}$ but not accumulatively). Note that for $\chi_1 = (123)$ and $\chi_2 = (121)$, where $\ell(\chi_1) = \ell(\chi_2) = 2$, $\hat{\mathbb{P}}^{\mu}_{\mathcal{C}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;\chi_1}\right)$ is close to $\hat{\mathbb{P}}^{\mu}_{\mathcal{C}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;\chi_2}\right)$ for each ℓ. In Fig. 13, the simulation shows that for each $\ell \in \{1, 2, \ldots, 20\}$, $\hat{\mathbb{P}}^{\mu}_{\mathcal{C}_n;\ell}\left(\mathcal{A}^{\mu}_{\mathcal{C}_n;\ell;\chi}\right)$ *decreases* as $\ell(\chi)$ *increases*, where $\chi \in \{(12),(123),(121),(1234)\}$.

Fig. 14 shows the variability between the recoverability probabilities for $\mathcal{G} =$

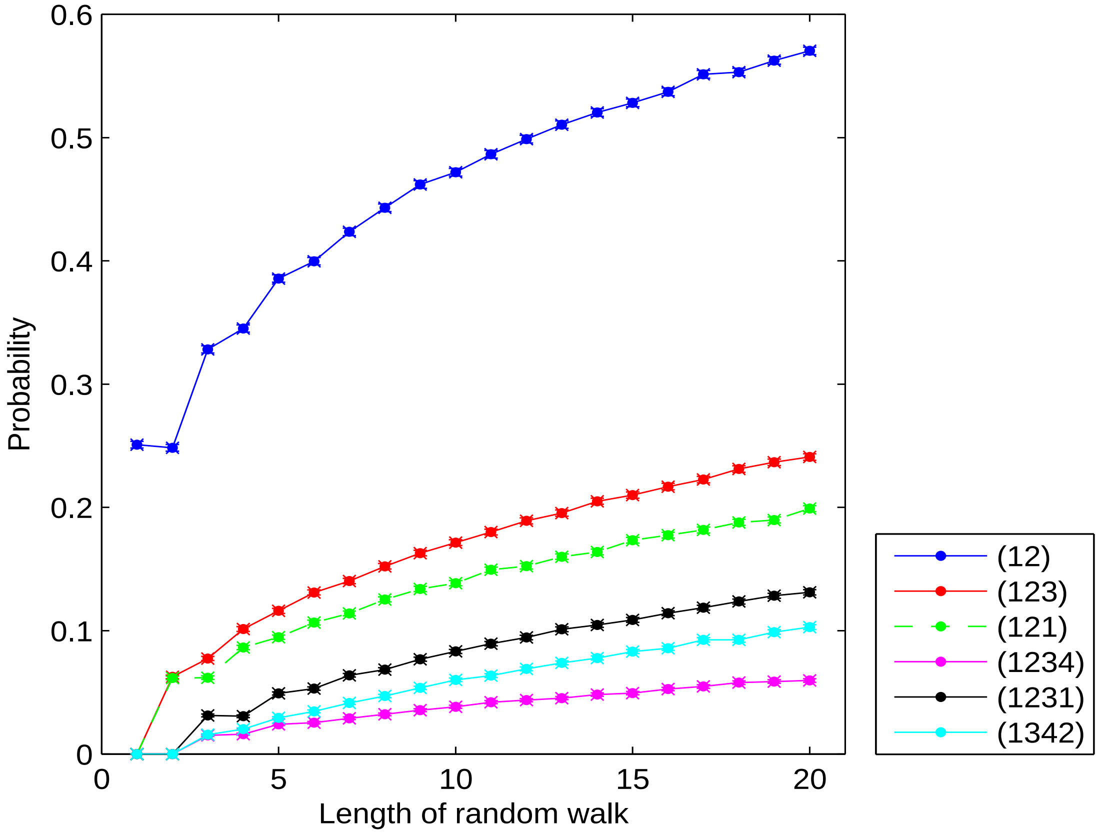


Figure 14: For $\mathcal{G} = \mathcal{G}(1000, 2)$, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ for different $\chi \in \{(12), (123), (121), (1234), (1231), (1342)\}$, where $\ell = 1, 2, \ldots, 20$. For each χ and for each ℓ, an estimate of a 95% confidence interval for $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ is also shown.

$\mathcal{G}(1000, 2)$. Although (123) and (121) have the same length, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;(123)}\right)$ is larger than $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;(121)}\right)$ for $\ell \geq 3$. Let us understand this behaviour by an example. We recall that for $\ell = 2$, $\mathbb{P}^{\mu}_{\mathcal{G};2}\left(\mathcal{A}^{\mu}_{\mathcal{G};2;(123)}\right) = \mathbb{P}^{\mu}_{\mathcal{G};2}\left(\mathcal{A}^{\mu}_{\mathcal{G};2;(121)}\right) = \frac{1}{16}$. For $\ell = 3$, (121) appears in the accumulative prime factorisation of these 4 possible realisations of $\mathcal{W}^{1}_{\mathcal{G};3}$: (1212),(1213),(121n),(121(n-1)). Thus $\mathbb{P}^{\mu}_{\mathcal{G};3}\left(\mathcal{A}^{\mu}_{\mathcal{G};3;(121)}\right) = \frac{4}{64} = \frac{1}{16}$. However for $\ell = 3$, (123) appears in the accumulative prime factorisation of these 5 possible realisation of $\mathcal{W}^{1}_{\mathcal{G};3}$: (1231),(1232),(1234),(1235),(1321). Not that for (1321), (123) does not appear in the accumulative prime factorisation of (132) but appears in the accumulative prime factorisation of (1321). Thus $\mathbb{P}^{\mu}_{\mathcal{G};3}\left(\mathcal{A}^{\mu}_{\mathcal{G};3;(123)}\right) = \frac{5}{64}$ for $\ell = 3$. We can also see the dependence of $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ on the selection of χ for $\ell(\chi) = 3$ in Fig. 14. As for the difference between the recoverability probabilities of (123) and (121), a similar explanation works for the difference between the recoverability probabilities of (1234),(1231) and (1342) for $\ell \geq 3$. However for $\mathcal{G} = \mathcal{G}(1000, 2)$, the simulation shows that for each $\ell \in \{1, 2, \ldots, 20\}$, $\hat{\mathbb{P}}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ *decreases* as $\ell(\chi)$ *increases*, where $\chi \in \{(12), (123), (121), (1234), (1231), (1342)\}$.

We now consider arbitrary connected r-regular graphs $\mathcal{G}_r$ on n vertices. We consider $2 < r < n - 1$. Let $n = 1,000$ and $m = 100,000$ (the sample size). For each ℓ and for each sample, we perform the following simulation steps.

1. We select an r-regular graph randomly using the MATLAB function createRandRegGraph and check if the selected graph is connected. If the selected graph is not connected, we re-sample a graph using createRandRegGraph until a connected r-regular graph is obtained.

2. For the sampled graph, we select a vertex μ with probability $\frac{1}{n}$ and select a walk w of fixed length $\ell_p \in \{1, 2, 3\}$ from those with μ as one of the endpoints with probability $\frac{1}{r^{\ell_p}}$. Then we check if w is a prime. If the selected w is not a prime, we re-select w with probability $\frac{1}{r^{\ell_p}}$ until a prime χ of length ℓ_p is obtained.

3. We perform a random walk of length ℓ starting from μ and check if χ appears in the accumulative prime factorisation of the realisation of the random walk.

Thus, we sample $m = 100,000$ connected r-regular graphs as in Section 9.1 and estimate the recoverability probability by using $\hat{p}_1(r, \ell, \ell_p) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{1}_{\mathcal{A}_i}$, where $1_{\mathcal{A}_i}$ returns 1 if in the i-th sample, the selected prime appears in the accumulative prime factorisation

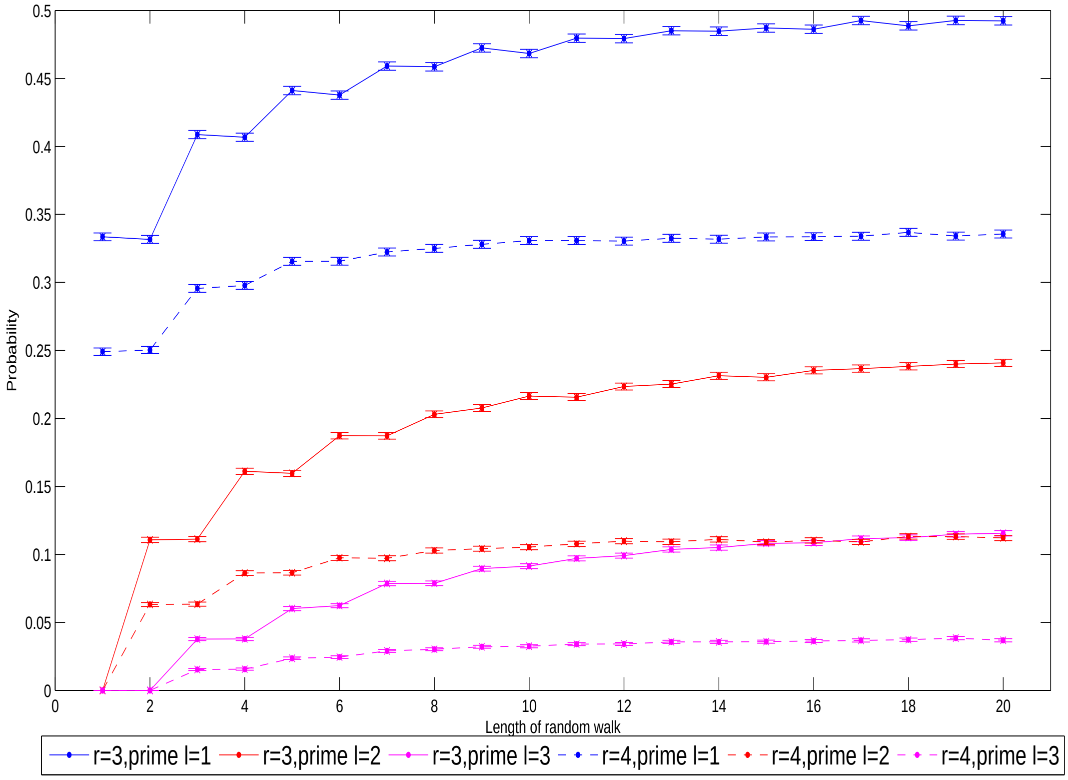


Figure 15: For $n = 1000$, $\hat{p}_1(r, \ell, \ell_p)$ where $\ell \in \{1, 2, \ldots, 20\}$ and $\ell_p \in \{1, 2, 3\}$. For each ℓ and ℓ_p, an estimate of a 95% confidence interval for the true recoverability probability is also shown.

of the random walk. In Fig. 15, the simulation shows that for $r = 3$ or $r = 4$, $\hat{p}_1(r, \ell, \ell')$ *decreases* as ℓ_p *increases* for each $\ell \in \{1, 2, \ldots, 20\}$ and for $\ell_p \in \{1, 2, 3\}$. One curiosity which is worth mentioning is that $\hat{p}_1(3, \ell, 3)$ seems to be close to $\hat{p}_1(4, \ell, 2)$ when $\ell \geq 15$ and sightly larger than $\hat{p}_1(4, \ell, 2)$ when $\ell = 19, 20$.

9.3 Open problems

9.3.1 Primes of length at least 2

In Section 9.2, we have seen some exploratory simulation results about the recoverability probability (based on the accumulative prime factorisation) of primes of length 1,2 or 3 on certain graphs. Let us consider a more fundamental question: what is the probability that a random walk on a connected r-regular graph $\mathcal{G}_r$ visits all the edges $\mathcal{E}(\chi)$ of a given prime χ in $\mathcal{G}_r$, where $\ell(\chi) \geq 2$? We recall that visiting all edges of a given prime χ is a necessary condition for recovering χ in the prime factorisation of a walk. Note that for $\chi = (\mu_1\mu_2\mu_1)$, where $\mu_1 \neq \mu_2$, that a walk visits all edges of χ means that the walk visits the edge $(\mu_1\mu_2)$. To avoid such cases, we assume that χ is *not* of the form $(\mu_1\mu_2\mu_1)$, where $\mu_1 \neq \mu_2$.

It is worth mentioning a result shown in [12], which obtains the probability of no visit to a small subset of vertices in a graph $\mathcal{G}$. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $|V| = n$. Let S_v be a subset of vertices in $\mathcal{G}$, where $|S_v| = o(n)$. To obtain the probability of no visit to S_v in $\mathcal{G}$, we form a new graph $\mathcal{H}$ from $\mathcal{G}$ by contracting S_v into a single vertex s_v and contracting the edges that are contained in S_v into loops. Lemma 11 in [12] shows that the probability of not visiting $S_v \subset \mathcal{V}(\mathcal{G})$ at time $t \geq T$, where T is the mixing time, can be obtained up to a multiplicative error $1 + O\left(\frac{1}{n^3}\right)$ from the probability of not visiting s_v in $\mathcal{H}$ at time t. This nice result and the technique of contracting a small set of vertices into a single vertex might be useful to find the probability that a random walk visits all edges of a given prime χ in $\mathcal{G}$. One can introduce an artificial vertex on each edge in the prime and contract the set of artificial vertices into a single vertex s_χ to form a new graph $\mathcal{H}$. Then the probability that a random walk on $\mathcal{G}$ visits all edges χ for the first time at time t ($t \geq T$) can be obtained from the probability that a random walk on $\mathcal{H}$ first visits s_χ at time t.

We recall that $\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}$ denotes the event that a prime χ appears in the accumulative prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$, and we factorise a walk accumulatively from the left end of the walk. Notice that $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$ is bounded above by the probability that $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$ visits all edges of χ. Let $\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell;\chi}$ denote the complementary event of $\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}$. Using the nature of the accumulative prime factorisation, we have

$$\begin{aligned}
&\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right) \\
&= \mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\mathcal{A}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)\mathbb{P}^{\mu}_{\mathcal{G};\ell-1}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell-1;\chi}\right) + \mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)\mathbb{P}^{\mu}_{\mathcal{G};\ell-1}\left(\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right) \\
&= \mathbb{P}^{\mu}_{\mathcal{G};\ell-1}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell-1;\chi}\right) + \mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)\left[1 - \mathbb{P}^{\mu}_{\mathcal{G};\ell-1}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)\right].
\end{aligned}$$

The last equation follows because if χ appears in the accumulative prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G};\ell-1}$, then χ appears in the accumulative prime factorisation of a realisation of $\mathcal{W}^{\mu}_{\mathcal{G};\ell}$. i.e. $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\mathcal{A}^{\mu}_{\mathcal{G};\ell-1;\chi}\right) = 1$. Thus, finding $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)$ is the key to obtain $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}\right)$. However, as we have seen in Section 9.2, $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)$ depends on the structure of $\mathcal{G}$. Making some assumption on the structure of $\mathcal{G}$ might be useful to find $\mathbb{P}^{\mu}_{\mathcal{G};\ell}\left(\mathcal{A}^{\mu}_{\mathcal{G};\ell;\chi}|\bar{\mathcal{A}}^{\mu}_{\mathcal{G};\ell-1;\chi}\right)$.

9.3.2 Number of primes recovered from a random walk

For a fixed regular graph $\mathcal{G}_r$, how many primes can be recovered from $\mathcal{W}^{\mu}_{\mathcal{G}_r;\ell}$? Since the motivation of this project is to obtain properties of network through primes of a network using random walks, it is essential to consider the number of primes visited by a random walk. One can use either the prime factorisation method in [19] or the accumulative prime factorisation method (factorise a walk accumulatively from the left end of the walk) to factorise a walk. The number of visited primes depends on the way to factorise a walk. The number of primes visited by a random walk might also be different from the number of *unique* primes visited by a random walk. One can consider the following questions:

1. Assume that $\mathcal{W}^{\mu}_{\mathcal{G}_r;\ell}$ ends at a vertex different from μ. What is the length of the simple path obtained by using the prime factorisation method in [19] likely to be?

2. How many unique simple paths can be obtained from the accumulative prime factorisation of the realisation of $\mathcal{W}^{\mu}_{\mathcal{G}_r;\ell}$?

3. Using either the prime factorisation method in [19] or the accumulative prime factorisation method, how many unique simple cycles can be obtained? What is the distribution of the length of the obtained simple cycles?

Ultimately, we are interested in getting all primes of a graph. What is the expected length of a random walk that recovers all the primes in a graph? This question is related to the notion of a *universal traversal sequence*. Let $\mathfrak{G}^{(r)}_{n;labeled}$ be the set of connected, edge-labeled r-regular graph. For a connected, edge-labeled r-regular graph, the edges incident to a vertex $\nu \in \mathcal{V}$ are labeled as $1, 2, \ldots, r$ in some way; it is not required that the labels be the same at both ends of an edge [1]. A sequence $s = (s_1, s_2, \ldots, s_\ell) \in \{1, 2, \ldots, r\}^\ell$ is a universal traversal sequence for $\mathfrak{G}^{(r)}_{n;labeled}$, if for *all* graphs $\mathcal{G}_r \in \mathfrak{G}^{(r)}_{n;labeled}$ and from *all* starting point, the walk defined by the sequence s visits each vertex of $\mathcal{G}_r$ at least once [1].

We give some previous works about universal traversal sequences but they are not meant to be exhaustive. Aleliunas *et al.* [2] used the probabilistic method to show that for $r \geq 2$, there exists a universal traversal sequence of length $O(r^2 n^3 \log n)$. Chandra *et al.* used a probabilistic argument similar to that in [2] to improve the upper bound for the length of a universal traversal sequence to $O(n^3 \log n)$ for $r \geq \lfloor \frac{n}{2} \rfloor$. Explicit constructions of universal traversal sequences of length $n^{4.03}$ for $r = 2$ were shown in [25] and of length $n^{O(\log n)}$ for $r = n - 1$ were shown in [24].

A natural variation is the universal *edge*-traversal sequence [5]. A sequence $s = (s_1, s_2, \ldots, s_\ell) \in \{1, 2, \ldots, r\}^\ell$ is a universal edge-traversal sequence for $\mathfrak{G}^{(r)}_{n;labeled}$, if for *all* graphs $\mathcal{G}_r \in \mathfrak{G}^{(r)}_{n;labeled}$ and from *all* starting point, the walk defined by the sequence s visits each edge of $\mathcal{G}_r$ at least once. It was shown in [5] that for nr even, if s is a universal edge-traversal sequence for $\mathfrak{G}^{(r)}_{n;labeled}$, then s is a universal edge-traversal sequence for $\mathfrak{G}^{(r)}_{n-r';labeled}$ for all $r' \geq r+1$. Unlike universal traversal sequences, research on universal edge-traversal sequences receives little attention.

We hope that a technique of proving that a sequence of a fixed length can be a universal traversal sequence, either by probabilistic methods or by explicit constructions, might be useful to find a sequence of a fixed length which recovers all the primes in a graph.

The open questions posed in this section suggest some possible directions for further research on the topic. The ideas provided with some questions are purely provisional and technicalities would have to be carefully verified.

References

[1] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, 2002. Recompiled in 2014 and available at http://www.stat.berkeley.edu/ aldous/RWG/book.html. Accessed in April 2014.

[2] R. Aleliunas, R. M. Karp, R. J. Lipton, L. Lovász and C. Rackoff, Random walks, universal traversal sequences, and the complexity of maze problems, *Proceedings of the 20th Annual Symposium on Foundations of Computer Science*, 218-223, 1979.

[3] A. Asztalos and Z. Toroczkai, Network discovery by generalized random walks, *Europhysics Letters* **92(5)**, 50008, 2010.

[4] N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, 2nd ed., 1993.

[5] A. Borodin, W. L. Ruzzo and M. Tompa, Lower bounds on the length of universal traversal sequences, *Journal of Computer and System Sciences* **45(2)**, 180-203., 1992.

[6] B. Bollobás, *Random Graphs*, Cambridge University Press, 2nd ed., 2001.

[7] R. A. Brualdi, *The Mutually Beneficial Relationship of Graphs and Matrices*, CBMS **115**, American Mathematical Society, 2011.

[8] S. Butler, *Eigenvalues and Structures of Graphs*, Ph.D. Dissertation, University of California, San Diego, 2008. Available at https://orion.math.iastate.edu/butler/PDF/dissertation.pdf.

[9] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky and P. Tiwari, The electrical resistance of a graph captures its commute and cover times, *Computational Complexity* **6(4)**, 312-340, 1997.

[10] C. Cooper and A. Frieze, The cover time of random regular graphs, *SIAM Journal on Discrete Mathematics* **18**, 728-740, 2005.

[11] C. Cooper and A. Frieze, The cover time of the giant component of $\mathcal{G}_{n,p}$, *Random Structures and Algorithms* **32**, 401-439, 2008.

[12] C. Cooper and A. Frieze, Component structure of the vacant set induced by a random walk on a random graph, *Random Structures and Algorithms* **42(2)**, 135-158, 2013.

[13] C. Cooper, T. Radzik and Y. Siantos, Estimating network parameters using random walks, *Proceedings of the 2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)*, 33-40, 2012.

[14] L. da F. Costa and G. Travieso, Exploring complex networks through random walks, *Physical Review E* **75(1)**, 16102, 2007.

[15] E. Estrada and J. A. Rodriguez-Veláquez, Subgraphs centrality in complex networks, *Physical Review E* **71**, 056103, 2005.

[16] E. Estrada, N. Hatano and M. Benzi, The physics of communicability in complex networks, *Physical Reports* **514**, 89-119, 2012.

[17] J. Friedman, *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*, Memoirs of the American Mathematical Society **195**, 2008.

[18] P.-L. Giscard, *A Graph Theoretic Approach to Matrix Functions and Quantum Dynamics*, Unpublished DPhil Thesis, University of Oxford, 2013.

[19] P.-L. Giscard, S. J. Thwaite and D. Jaksch, Continued fractions and unique factorization on digraphs, under review *arXiv:1202.5523*, 2012.

[20] P.-L. Giscard, S. J. Thwaite and D. Jaksch, Evaluating matrix functions by resummations on graphs: the method of path-sums, *SIAM Journal on Matrix Analysis & Applications* **34(2)**, 445-469, 2013.

[21] F. Harary, *Graph Theory*, Addison-Wesley, Reading, Massachusetts, 1969.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 2nd ed., 2013.

[23] E. I. Jury, *Theory and Application of the Z-Transform Method*, New York; London: John Wiley and Sons, 1964.

[24] H. J. Karloff, R. Paturi and J. Simon, Universal traversal sequences of length $n^{O(\log n)}$ for cliques, *Information Processing Letters* **28(5)**, 241-243, 1988.

[25] M. Koucký, Log-space constructible universal traversal sequences for cycles of length $O(n^{4.03})$, *Theoretical Computer Science* **296(1)**, 117-144, 2003.

[26] E. Kreyszig, *Introductory Functional Analysis with Applications*, New York; London: John Wiley and Sons, 1978.

[27] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.

[28] B. Liu and G. Li, A note on the largest eigenvalue of non-regular graphs, *Electronic Journal of Linear Algebra* **17**, 54-61, 2008.

[29] B. Liu, Y. Huang and Z. You, On λ_1-extremal non-regular graphs, *Electronic Journal of Linear Algebra* **18**, 735-744, 2009.

[30] B. Liu, L. Mu-Huo and Z. You, Erratum to 'A note on the largest eigenvalue of non-regular graphs', *Electronic Journal of Linear Algebra* **18**, 64-68, 2009.

[31] B. Liu, J. Shen and X. Wang, On the largest eigenvalue of non-regular graphs, *Journal of Combinatorial Theory, Series B* **97**, 1010-1018, 2007.

[32] L. Lovász, Random walks on graphs: A survey, *Combinatorics, Paul Erdös is Eighty* **2**, 1-46, 1993.

[33] M. E. J. Newman, *Networks: an introduction*, Oxford University Press, 2010.

[34] H. A. Priestley, *Introduction to Complex Analysis*, Oxford University Press, 2nd ed., 2003.

[35] N. J. A. Sloane, Sequence A002851, *The On-Line Encyclopedia of Integers Sequences*, published electronically at http://oeis.org, 2014.

[36] J. Shu and Y. Wu, Sharp upper bounds on the spectral radius of graphs, *Linear Algebra and its Applications* **377**, 241-248, 2004.

[37] X.-D. Zhang, Eigenvectors and eigenvalues of non-regular graphs, *Linear Algebra and its Applications* **409**, 79-86, 2005.

Appendix A

A.1 Blockwise inversion formula

We explain how the blockwise inversion of an $n \times n$ matrix B works. We can write B as a 2×2 block matrix

$$B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},$$

where B_1 is a $k \times k$ matrix, $1 \leq k \leq n-1$, and B_4 is an $(n-k) \times (n-k)$ matrix. The well-known blockwise inversion formula for a 2×2 block matrix is given by [22]

$$B^{-1} = \begin{pmatrix} (B_1 - B_2 B_4^{-1} B_3)^{-1} & -(B_1 - B_2 B_4^{-1} B_3)^{-1} B_2 B_4^{-1} \\ -B_4^{-1} B_3 (B_1 - B_2 B_4^{-1} B_3)^{-1} & (B_4 - B_3 B_1^{-1} B_2)^{-1} \end{pmatrix}, \tag{A.1}$$

assuming all the relevant inverses exist.

Notice that

$$\begin{aligned} B &= \begin{pmatrix} B_1 & 0 \\ B_3 & I_{n-k} \end{pmatrix} \begin{pmatrix} I_k & B_1^{-1} B_2 \\ 0 & B_4 - B_3 B_1^{-1} B_2 \end{pmatrix}, \quad \text{if } B_1 \text{ is invertible} && \text{(A.2)} \\ &= \begin{pmatrix} I_k & B_2 \\ 0 & B_4 \end{pmatrix} \begin{pmatrix} B_1 - B_2 B_4^{-1} B_3 & 0 \\ B_4^{-1} B_3 & I_{n-k} \end{pmatrix}, \quad \text{if } B_4 \text{ is invertible.} && \text{(A.3)} \end{aligned}$$

Assume that B is invertible, then $\det(B) \neq 0$. If in addition B_1 is invertible, we have $\det(B) = \det(B_1)\det(B_4 - B_3 B_1^{-1} B_2) \neq 0$ by Eq. (A.2), then $B_4 - B_3 B_1^{-1} B_2$ is invertible. If B_4 is invertible, we have $\det(B) = \det(B_4)\det(B_1 - B_2B_4^{-1}B_3) \neq 0$ by Eq. (A.3), then $B_1 - B_2B_4^{-1}B_3$ is invertible. Thus if B, B_1 and B_4 are invertible, we can use the blockwise inversion formula in Eq. (A.1) to obtain the inverse of B.

A.2 Application to walk generating matrix

We can write $A_{\mathcal{G}}$ and $I - zA_{\mathcal{G}}$ in the form of 2×2 block matrices

$$A_{\mathcal{G}} = \begin{pmatrix} A_{11;\mathcal{G}} & A_{12;\mathcal{G}} \\ A_{21;\mathcal{G}} & A_{22;\mathcal{G}} \end{pmatrix}, \quad I - zA_{\mathcal{G}} = \begin{pmatrix} I_k - zA_{11;\mathcal{G}} & -zA_{12;\mathcal{G}} \\ -zA_{21;\mathcal{G}} & I_{n-k} - zA_{22;\mathcal{G}} \end{pmatrix},$$

where $A_{11;\mathcal{G}}$ is a $k \times k$ matrix, $1 \leq k \leq n-1$, and $A_{22;\mathcal{G}}$ is an $(n-k) \times (n-k)$ matrix. Note that $A_{12;\mathcal{G}} = (A_{21;\mathcal{G}})^T$ as $A_{\mathcal{G}}$ is symmetric. We know that $I - zA_{\mathcal{G}}$ is invertible when $|z| < \frac{1}{||A_{\mathcal{G}}||}$. Moreover, $I_k - zA_{11;\mathcal{G}}$ and $I_{n-k} - zA_{22;\mathcal{G}}$ are invertible when $|z| \leq \frac{1}{||A_{11;\mathcal{G}}||} \wedge \frac{1}{||A_{22;\mathcal{G}}||}$. Thus, using the blockwise inversion formula given in Eq. (A.1), $(I - zA_{\mathcal{G}})^{-1}$ can be expressed in the following form,

$$\begin{pmatrix} I_k - zA_{11;\mathcal{G}} & -zA_{12;\mathcal{G}} \\ -zA_{21;\mathcal{G}} & I_{n-k} - zA_{22;\mathcal{G}} \end{pmatrix}^{-1} = \begin{pmatrix} G_{11;\mathcal{G}}(z) & [G_{21;\mathcal{G}}(z)]^T \\ G_{21;\mathcal{G}}(z) & G_{22;\mathcal{G}}(z) \end{pmatrix}, \tag{A.4}$$

where

$$G_{11;\mathcal{G}}(z) = [I_k - zA_{11;\mathcal{G}} - z^2A_{12;\mathcal{G}}(I_{n-k} - zA_{22;\mathcal{G}})^{-1}A_{21;\mathcal{G}}]^{-1}, \tag{A.5}$$

$$G_{21;\mathcal{G}}(z) = z(I_{n-k} - zA_{22;\mathcal{G}})^{-1}A_{21;\mathcal{G}}G_{11;\mathcal{G}}(z), \tag{A.6}$$

$$G_{22;\mathcal{G}}(z) = [I_{n-k} - zA_{22;\mathcal{G}} - z^2A_{21;\mathcal{G}}(I_k - zA_{11;\mathcal{G}})^{-1}A_{12;\mathcal{G}}]^{-1}. \tag{A.7}$$

Appendix B

B.1 Proof of Theorem 3 using walk generating functions

Let the first row of $A_{\mathcal{G}}$ correspond to ν_1 and the second row of $A_{\mathcal{G}}$ correspond to ν_2 without loss of generality, i.e. when we say the (ν_1,ν_2) entry of $A_{\mathcal{G}}$, we mean the $(1,2)$ entry of $A_{\mathcal{G}}$. Firstly we obtain $m(\mathfrak{W}^{\mu\nu}_{\mathcal{K}_n\setminus\{e\};\ell})$ for $\ell \geq 1$, where $\nu \in \mathcal{V}$. For simplicity, we denote $\mathcal{K}_n \setminus \{e\}$ by $\mathcal{G}_1$. Notice that $(A_{\mathcal{G}_1})_{12} = (A_{\mathcal{G}_1})_{21} = 0$ and $A_{22;\mathcal{G}_1}$ is the adjacency matrix of $\mathcal{K}_{n-2}$. Thus $(I_2 - zA_{11;G_1}) = I_2$ and $(I_{n-2} - zA_{22;\mathcal{G}_1})$ are invertible when $|z| < \frac{1}{||A_{\mathcal{G}_1}||} \leq \frac{1}{||A_{22;\mathcal{G}_1}||}$. By letting $\mathcal{G} = \mathcal{G}_1$ and $k = 2$ in Eq. (A.4) and by Eq. (A.5, A.6, A.7), we have

$$
\begin{aligned}
G_{11;\mathcal{G}_1}(z) &= \left[I - z\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} - z^2 A_{12;\mathcal{G}_1}(I_{n-2} - zA_{22;\mathcal{G}_1})^{-1}(A_{12;\mathcal{G}_1})^T \right]^{-1}, \\
G_{21;\mathcal{G}_1}(z) &= z(I_{n-2} - zA_{22;\mathcal{G}_1})^{-1}(A_{12;\mathcal{G}_1})^T G_{11;\mathcal{G}_1}(z), \\
G_{22;\mathcal{G}_1}(z) &= \left[I_{n-k} - zA_{22;\mathcal{G}} - z^2 (A_{12;\mathcal{G}_1})^T \left[I - z\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right]^{-1} A_{12;\mathcal{G}_1} \right]^{-1}.
\end{aligned}
$$

We recall that $(I_{n-2} - zA_{22;\mathcal{G}_1})^{-1}$ can be written as $\sum_{\ell=0}^{\infty}(zA_{22;\mathcal{G}_1})^{\ell}$ when $|z| < \frac{1}{||A_{\mathcal{G}_1}||} \leq \frac{1}{||A_{22;\mathcal{G}_1}||}$. Let a_{ℓ}^{diag} and a_{ℓ}^{off} denote the diagonal entry and the off-diagonal entry of $A^{\ell}_{22;\mathcal{G}_1}$ respectively. Note that

$$
\begin{aligned}
a_{\ell}^{diag} &= (n-3)a_{\ell-1}^{off}, \\
a_{\ell}^{off} &= (n-4)a_{\ell-1}^{off} + a_{\ell-1}^{diag},
\end{aligned}
$$

with $a_0^{diag} = 1$, $a_0^{off} = 0$ and $a_1^{off} = 1$. Solving the recurrence relations, we get $a_\ell^{diag} = \frac{n-3}{n-2}[(n-3)^{\ell-1} - (-1)^{\ell-1}]$ and $a_\ell^{off} = \frac{1}{n-2}[(n-3)^\ell - (-1)^\ell]$. Thus for $|z| < \frac{1}{||A_{\mathcal{G}_1}||}$, the diagonal entries of $(I_{n-2} - zA_{22;\mathcal{G}_1})^{-1}$ are $1 + \sum_{\ell=1}^{\infty} \left\{\frac{n-3}{n-2}[(n-3)^{\ell-1} - (-1)^{\ell-1}]z^\ell\right\} = \frac{1-(n-4)z}{[1-(n-3)z](1+z)}$ and its off-diagonal entries are $\sum_{\ell=0}^{\infty} \left\{\frac{1}{n-2}[(n-3)^\ell - (-1)^\ell]z^\ell\right\} = \frac{z}{[1-(n-3)z](1+z)}$. Note also that each entry of $A_{12;\mathcal{G}_1}$ is 1. Then we have,

$$G_{11;\mathcal{G}_1}(z) = \begin{pmatrix} \frac{1-(n-3)z-(n-2)z^2}{1-(n-3)z-2(n-2)z^2} & \frac{(n-2)z^2}{1-(n-3)z-2(n-2)z^2} \\ \frac{(n-2)z^2}{1-(n-3)z-2(n-2)z^2} & \frac{1-(n-3)z-(n-2)z^2}{1-(n-3)z-2(n-2)z^2} \end{pmatrix},$$

$$G_{11;\mathcal{G}_1}(1/z) = \begin{pmatrix} \frac{z^2-(n-3)z-(n-2)}{z^2-(n-3)z-2(n-2)} & \frac{(n-2)}{z^2-(n-3)z-2(n-2)} \\ \frac{(n-2)}{z^2-(n-3)z-2(n-2)} & \frac{z^2-(n-3)z-(n-2)}{z^2-(n-3)z-2(n-2)} \end{pmatrix},$$

$$G_{21;\mathcal{G}_1}(z) = \frac{z}{1-(n-3)z-2(n-2)z^2} \begin{pmatrix} 1 & \dots & 1 \\ 1 & \dots & 1 \end{pmatrix}^T,$$

$$G_{21;\mathcal{G}_1}(1/z) = \frac{z}{z^2-(n-3)z-2(n-2)} \begin{pmatrix} 1 & \dots & 1 \\ 1 & \dots & 1 \end{pmatrix}^T.$$

The poles of each entry of $G_{11;\mathcal{G}_1}(1/z)$ and $G_{12;\mathcal{G}_1}(1/z) = (G_{12;\mathcal{G}_1}(1/z))^T$ are then $\frac{n-3+c(n)}{2}$ and $\frac{n-3-c(n)}{2}$, where $c(n) = \sqrt{n^2+2n-7}$. Following Eq. (9), we have

$$\operatorname{Res}_{z_0=\frac{n-3+c(n)}{2}} \left[(G_{11;\mathcal{G}_1}(1/z))_{\nu_1\nu_1}\right] z^{\ell-1} = \frac{(n-2)[n-3+c(n)]^{\ell-1}}{2^{\ell-1}c(n)},$$
$$\operatorname{Res}_{z_0=\frac{n-3-c(n)}{2}} \left[(G_{11;\mathcal{G}_1}(1/z))_{\nu_1\nu_1}\right] z^{\ell-1} = -\frac{(n-2)[n-3-c(n)]^{\ell-1}}{2^{\ell-1}c(n)}.$$

Then by Eq. (8), $m\left(\mathfrak{W}_{\mathcal{G}_1;\ell}^{\nu_1\nu_1}\right)$ is given by

$$\mathcal{Z}^{-1}\left[(G_{11;\mathcal{G}_1}(1/z))_{\nu_1\nu_1}\right](\ell) = \frac{(n-2)[n-3+c(n)]^{\ell-1}}{2^{\ell-1}c(n)} - \frac{(n-2)[n-3-c(n)]^{\ell-1}}{2^{\ell-1}c(n)}.$$

Similarly, we obtain $m\left(\mathfrak{W}^{\nu_1\nu_2}_{\mathcal{G}_1;\ell}\right)$ and $m\left(\mathfrak{W}^{\nu_1\nu}_{\mathcal{G}_1;\ell}\right)$ where $\nu \in \mathcal{V} \setminus \{\nu_1, \nu_2\}$, which are given by

$$\mathcal{Z}^{-1}\left[\left(G_{11;\mathcal{G}_1}(1/z)\right)_{\nu_1\nu_2}\right](\ell) = \frac{(n-2)[n-3+c(n)]^{\ell-1}}{2^{\ell-1}c(n)} - \frac{(n-2)[n-3-c(n)]^{\ell-1}}{2^{\ell-1}c(n)},$$
$$\mathcal{Z}^{-1}\left[\left(G_{12;\mathcal{G}_1}(1/z)\right)_{\nu_1\nu}\right](\ell) = \frac{[n-3+c(n)]^{\ell}}{2^{\ell}c(n)} - \frac{[n-3-c(n)]^{\ell}}{2^{\ell}c(n)}.$$

Then we get $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{G}_1;\ell}\right) = \sum_{\nu'\in\mathcal{V}} m\left(\mathfrak{W}^{\nu_1\nu'}_{\mathcal{G}_1;\ell}\right)$, which is given by

$$\begin{aligned}
&m\left(\mathfrak{W}^{\nu_1}_{\mathcal{G}_1;\ell}\right)\\
&= \mathcal{Z}^{-1}\left[\left(G_{11;\mathcal{G}_1}(1/z)\right)_{\nu_1\nu_1}\right](\ell) + \mathcal{Z}^{-1}\left[\left(G_{11;\mathcal{G}_1}(1/z)\right)_{\nu_1\nu_2}\right](\ell) +\\
&\quad (n-2)\mathcal{Z}^{-1}\left[\left(G_{11;\mathcal{G}_1}(1/z)\right)_{\nu_1\nu}\right](\ell)\\
&= \frac{[n-1+c(n)][n-3+c(n)]^{\ell} - [n-1-c(n)][n-3-c(n)]^{\ell}}{2^{\ell+1}c(n)}.
\end{aligned}$$

Knowing that $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n;\ell}\right) = (n-1)^{\ell}$, by Eq. (3) we get

$$\begin{aligned}
\mathbb{P}^{\nu_1}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\nu_1}_{\mathcal{K}_n;\ell;e}) &= 1 - \frac{m\left(\mathfrak{W}^{\nu_1}_{\mathcal{G}_1;\ell}\right)}{m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n;\ell}\right)}\\
&= 1 - \frac{[n-1+c(n)][n-3+c(n)]^{\ell} - [n-1-c(n)][n-3-c(n)]^{\ell}}{2^{\ell+1}(n-1)^{\ell}c(n)}.
\end{aligned}$$

B.2 Proof of Theorem 3 using recurrence relations

We have the following recurrence relations,

$$m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = \sum_{\alpha\in\mathcal{V}\setminus\{\nu_1,\nu_2\}} m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) \tag{B.1}$$

$$m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = \sum_{\nu\in\mathcal{V}} m\left(\mathfrak{W}^{\nu}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) - m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right), \quad \text{where } \alpha\in\mathcal{V}\setminus\{\nu_1,\nu_2\} \tag{B.2}$$

$$m\left(\mathfrak{W}^{\nu_2}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = \sum_{\alpha\in\mathcal{V}\setminus\{\nu_1,\nu_2\}} m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right). \tag{B.3}$$

By symmetry, we have $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = m\left(\mathfrak{W}^{\nu_2}_{\mathcal{K}_n\setminus\{e\};\ell}\right)$, and $m\left(\mathfrak{W}^{\alpha_1}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) = m\left(\mathfrak{W}^{\alpha_2}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right)$ for any $\alpha_1,\alpha_2\in\mathcal{V}\setminus\{\nu_1,\nu_2\}$ and $\alpha_1\neq\alpha_2$. We have the initial conditions that $m\left(\mathfrak{W}^{\nu}_{\mathcal{K}_n\setminus\{e\};0}\right) = 1$ for $\nu\in\mathcal{V}$, and $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};1}\right) = n-2$ and $m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};1}\right) = n-1$ for $\alpha\in\mathcal{V}\setminus\{\nu_1,\nu_2\}$.

We can simplify (B.1) - (B.3) into

$$m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = (n-2)m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right), \quad \text{where } \alpha\in\mathcal{V}\setminus\{\nu_1,\nu_2\} \tag{B.4}$$

$$m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = 2m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) + (n-3)m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) \tag{B.5}$$

Substituting (B.5) into (B.4), we get

$$\begin{aligned}
m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) &= 2(n-2)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-2}\right) + (n-2)(n-3)m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-2}\right) \\
&= 2(n-2)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-2}\right) + 2(n-2)(n-3)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-3}\right) + \\
&\quad (n-2)(n-3)^3 m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell-3}\right) \\
&= \vdots
\end{aligned}$$

$$= 2(n-2)\sum_{j=2}^{\ell-1}(n-3)^{j-2}m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-j}\right) + (n-2)(n-3)^{\ell-2}m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};1}\right)$$
$$= 2(n-2)\sum_{j=2}^{\ell-1}(n-3)^{j-2}m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-j}\right) + (n-1)(n-2)(n-3)^{\ell-2}.$$

Thus $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) = 2(n-2)\sum_{j=3}^{\ell-1}(n-3)^{j-3}m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-j}\right) + (n-1)(n-2)(n-3)^{\ell-3}$ and

$$\begin{aligned}
&m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) - m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right)\\
&= 2(n-2)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-2}\right) + 2(n-2)\sum_{j=3}^{\ell-1}(n-3)^{j-3}(n-4)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-j}\right) +\\
&\quad (n-1)(n-2)(n-3)^{\ell-3}(n-4).
\end{aligned}$$

From above, we derive $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = (n-3)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-1}\right) + 2(n-2)m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell-2}\right)$, which is a linear homogeneous recurrence relation.

We use the ansatz that the solution of the linear homogeneous relation is of the form $a^\ell = (n-3)a^{\ell-1} + 2(n-4)a^{\ell-2}$. Solving for a, we obtain $a_1 = \frac{n-3+c(n)}{2}$ and $a_2 = \frac{n-3-c(n)}{2}$, where $c(n) = \sqrt{n^2-2n-7}$. Notice that $c_1a_1^\ell + c_2a_2^\ell$, where $c_1, c_2 \in \mathbb{R}$, also satisfies the recurrence relation. We use the initial conditions that $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};0}\right) = 1$, $m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};1}\right) = n-2$ to solve for c_1 and c_2. We obtain $c_1 = \frac{n-1+c(n)}{2c(n)}$ and $c_2 = -\frac{n-1-c(n)}{2c(n)}$. Thus

$$m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = \frac{[n-1+c(n)][n-3+c(n)]^\ell - [n-1-c(n)][n-3-c(n)]^\ell}{2^{\ell+1}c(n)}$$
$$m\left(\mathfrak{W}^{\alpha}_{\mathcal{K}_n\setminus\{e\};\ell}\right) = \frac{[n-1+c(n)][n-3+c(n)]^{\ell+1} - [n-1-c(n)][n-3-c(n)]^{\ell+1}}{2^{\ell+2}(n-2)c(n)}.$$

Then it follows that

$$
\begin{aligned}
\mathbb{P}^{\nu_1}_{\mathcal{K}_n;\ell}(\mathcal{A}^{\nu_1}_{\mathcal{K}_n;\ell;e}) &= 1 - \frac{m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n\setminus\{e\};\ell;\ell}\right)}{m\left(\mathfrak{W}^{\nu_1}_{\mathcal{K}_n;\ell}\right)} \\
&= 1 - \frac{[n-1+c(n)][n-3+c(n)]^\ell - [n-1-c(n)][n-3-c(n)]^\ell}{2^{\ell+1}(n-1)^\ell c(n)}.
\end{aligned}
$$

B.3 Proof of Theorem 4

Notice that $\mathcal{G}_1 = \mathcal{C}_n \setminus \{e\}$ is a path graph with n vertices, and ν_1 and ν_2 are the endpoints of the path graph. We denote $\nu_{(d)}$ as the vertex which is d steps away from ν_1, with $\nu_{(0)} = \nu_1$ and $\nu_{(n-1)} = \nu_2$. We use a result from [19] that the walk generating function $g^{\nu_1\nu_{(d)}}_{\mathcal{G}_1}(z)$ is given by

$$
g^{\nu_1\nu_{(d)}}_{\mathcal{G}_1}(z) = \frac{z^d Q_{n-d-1}(z)}{Q_n(z)},
$$

where $|z| < \frac{1}{||A_{\mathcal{G}_1}||}$ and $Q_x(z) = {}_2F_1[\frac{1}{2} - \frac{x}{2}, -\frac{x}{2}; -x; 4z^2]$ is the Gauss hypergeometric function.

We recall that the denominator of $g^{\nu_1\nu_{(d)}}_{\mathcal{G}_1}(z)$ is the characteristic polynomial of $\mathcal{G}_1$ evaluated at $\frac{1}{z}$. Knowing that the eigenvalues of $\mathcal{G}_1$ are $\left\{\lambda^{(n)}_j = 2\cos\left(\frac{\pi j}{n+1}\right)\right\}_{j=1}^{n}$, we have

$$
Q_n(z) = \prod_{j=1}^{n}(1 - z\lambda_j),
$$

and thus we get

$$
g^{\nu_1\nu_{(d)}}_{\mathcal{G}_1}(z) = \frac{z^d \prod_{j=1}^{n-d-1}\left(1 - z\lambda^{(n-d-1)}_j\right)}{\prod_{k=1}^{n}\left(1 - z\lambda^{(n)}_k\right)}, \quad \text{for } |z| < \frac{1}{||A_{\mathcal{G}_1}||},
$$

$$g_{\mathcal{G}_1}^{\nu_1\nu_{(d)}}(1/z) = \frac{z\prod_{j=1}^{n-d-1}\left(z-\lambda_j^{(n-d-1)}\right)}{\prod_{k=1}^{n}\left(z-\lambda_k^{(n)}\right)}, \quad \text{for } |z| > ||A_{\mathcal{G}_1}||.$$

From the last expression, we can see that the poles of $g_{\mathcal{G}_1}^{\nu_1\nu_{(d)}}(1/z)$ are the eigenvalues of $\mathcal{G}_1$, each with multiplicity 1. By Eq. (9), for $1 \le k \le n$,

$$\mathrm{Res}_{z_0=\lambda_k^n}\left[g_{\mathcal{G}_1}^{\nu_1\nu_{(d)}}(1/z)\right] z^{\ell-1} = \frac{\left(\lambda_k^{(n)}\right)^{\ell}\prod_{j=1}^{n-d-1}\left(\lambda_k^{(n)}-\lambda_j^{(n-d-1)}\right)}{\prod_{h=1;h\neq k}^{n}\left(\lambda_k^{(n)}-\lambda_h^{(n)}\right)}.$$

Then by Eq. (8), $m\left(\mathfrak{W}_{\mathcal{G}_1;\ell}^{\nu_1\nu_{(d)}}\right)$ is given by

$$\begin{aligned}\mathcal{Z}^{-1}\left[g_{\mathcal{G}_1}^{\nu_1\nu_{(d)}}(1/z)\right] &= \sum_{k=1}^{n}\frac{\left(\lambda_k^{(n)}\right)^{\ell}\prod_{j=1}^{n-d-1}\left(\lambda_k^{(n)}-\lambda_j^{(n-d-1)}\right)}{\prod_{h=1;h\neq k}^{n}\left(\lambda_k^{(n)}-\lambda_h^{(n)}\right)} \\ &= \sum_{k=1}^{n}\frac{2^{\ell-d}\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell}\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi j}{n-d}\right)\right]}{\prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi h}{n+1}\right)\right]},\end{aligned}$$

and thus we get $m\left(\mathfrak{W}_{\mathcal{G}_1;\ell}^{\nu_1}\right) = \sum_{d=0}^{n-1} m\left(\mathfrak{W}_{\mathcal{G}_1;\ell}^{\nu_1\nu_{(d)}}\right)$, which is given by

$$m\left(\mathfrak{W}_{\mathcal{G}_1;\ell}^{\nu_1}\right) = \sum_{d=0}^{n-1}\sum_{k=1}^{n}\frac{2^{\ell-d}\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell}\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi j}{n-d}\right)\right]}{\prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi h}{n+1}\right)\right]},$$

where we define $\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi j}{n-d}\right)\right] = 1$ when $d = n-1$.

Knowing that $m\left(\mathfrak{W}_{\mathcal{C}_n;\ell}^{\nu_1}\right) = 2^{\ell}$, by Eq. (3) we get

$$\mathbb{P}_{\mathcal{C}_n;\ell}^{\nu_1}(\mathcal{A}_{\mathcal{C}_n;\ell;e}^{\nu_1}) = 1 - \sum_{d=0}^{n-1}\sum_{k=1}^{n}\frac{\left[\cos\left(\frac{\pi k}{n+1}\right)\right]^{\ell}\prod_{j=1}^{n-d-1}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi j}{n-d}\right)\right]}{2^{d}\prod_{h=1;h\neq k}^{n}\left[\cos\left(\frac{\pi k}{n+1}\right)-\cos\left(\frac{\pi h}{n+1}\right)\right]}.$$

Appendix C Proof of $\rho\left(A_{\mathcal{G}_r \setminus \{e\}}\right) < \rho\left(A_{\mathcal{G}_r}\right)$

Let $A \in \mathcal{M}_n$ be an irreducible and nonnegative matrix. We show that $A + B$ is irreducible whenever $B \in \mathcal{M}_n$ is nonnegative, and $\rho(A + B) > \rho(A)$ whenever B is nonnegative and $B \neq 0$.

To prove the first part, we proceed by contradiction. Assume that $A+B$ is reducible. Then by Definition 1, there exists a permutation matrix P such that

$$P^T(A+B)P = P^TAP + P^TBP = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix},$$

where $A_1 \in \mathcal{M}_r$ and $A_3 \in \mathcal{M}_{n-r}$, for some r where $1 \leq r \leq n-1$. As A is irreducible, P^TAP is not of a form as $\begin{pmatrix} A_1' & A_2' \\ 0 & A_3' \end{pmatrix}$, where $A_1' \in \mathcal{M}_{r'}$ and $A_3' \in \mathcal{M}_{n-r'}$, for some r' where $1 \leq r' \leq n-1$. Moreover B is nonnegative, thus P^TBP is nonnegative. Hence $A + B$ must be irreducible.

Let $C \in \mathcal{M}_n$ and $(|A|)_{ij} \equiv |(A)_{ij}|$ (entrywise absolute value) for $i, j = 1, \ldots, n$. If $|A| \leq C$, then $\rho(A) \leq \rho(|A|) \leq \rho(C)$ [22](Theorem 8.1.18). Thus, we have $\rho(A+B) \geq \rho(A)$. If $\rho(A+B) = \rho(A)$, then there exists some $x \neq 0$ such that $Ax = \lambda x$ with $|\lambda| = \rho(A+B) = \rho(A)$, and $\rho(A+B)|x| = |\lambda x| = |Ax| \leq |A||x| \leq (A+B)|x|$ $(*)$.

As $A + B$ is irreducible and nonnegative, so is $(A+B)^T$. Hence by the Perron-Frobenius Theorem (Theorem 5), $(A+B)^T$ has a positive eigenvector y such that $(A+B)^Ty = \rho(A+B)y$. This gives $(A+B)|x| = \rho(A+B)|x|$ since $y^T\left[(A+B)|x| - \rho(A+B)|x|\right] = \rho(A+B)y^T|x| - \rho(A+B)y^T|x| = 0$.

Given $(*)$ and $(A+B)|x| = \rho(A+B)|x|$, we obtain $|Ax| = A|x| = (A+B)|x|$. As $B \geq 0$ and $x \neq 0$, B must be a zero matrix. Thus $\rho(A+B) > \rho(A)$ whenever B is nonnegative and $B \neq 0$.