Yale University

Preprint

Learning Random Regular Graphs

Dongqu Chen
Department of Computer Science
Yale University

YALEU/DCS/TR-1518
September 2015

Learning Random Regular Graphs

Dongqu Chen
Department of Computer Science
Yale University
dongqu.chen@yale.edu

Abstract

The family of random regular graphs is a classic topic in the realms of graph theory, combinatorics and computer science. In this paper we study the problem of learning random regular graphs from random paths. A random regular graph is generated uniformly at random and in a standard label-guided graph exploration setting, the edges incident from a node in the graph have distinct local labels. The input data to the statistical query oracle are path-vertex pairs (x, v) where x is a random uniform path (a random sequence of edge labels) and v is the vertex of the graph reached on the path x starting from a particular start vertex v_0. We present a comprehensive study and prove positive results on the convergence of random walks on many types of random regular graphs. In addition to the theoretical results, we generalize Angluin and Chen's learning algorithm to learning random regular graphs from uniform paths in the statistical query model. Extensive experiments demonstrate the efficiency and accuracy of the algorithm.

1 Introduction

Random walks on graphs have long served as a fundamental topic in the study of Markov chains and also as an important tool in machine learning research. On the other hand, regular graphs are widely studied in theoretical computer science for their important role in computational graph models and their applications. As most strong properties usually don't hold for all regular graphs, it is natural to ask whether we can pursue positive results for "almost all" regular graphs. This is addressed by studying high-probability properties of uniformly generated random regular graphs. In recent decades random regular graphs have gathered more and more attention in computer science, combinatorics and graph theory (see the related works section). Nevertheless, the study of random walks on random regular graphs is relatively limited. This paper aims to fill this gap with a comprehensive study of the varieties of random regular graphs listed in Table 1. Detailed definitions of the random graph models are provided in Section 2.1. The notations in Table 1 are used throughout this paper. Our main contributions are the positive results on the fast convergence of random walks on random regular graphs, which fill the gap in the research on random regular graphs. In addition to the theoretical results, we are able to generalize Angluin and Chen's algorithm [1] to learning random regular graphs (i.e., almost all regular graphs) from random paths in the statistical query model.

Random out-regular multigraphs ($\mathrm{RMG}^+(s)$) are the most well-studied among the family of random regular graphs, mainly because the freedom and independence of the edge selections makes the analysis simple and direct. This is also due to the important role of deterministic finite automaton (DFA) in computer science, as the underlying automaton graph of a random DFA is exactly a $\mathrm{RMG}^+(s)$ ([1, 15, 27]). In the context of DFA learning, Angluin and Chen [1] have proved the fast convergence of random walks on a $\mathrm{RMG}^+(s)$. In our paper, we first start with the slightly more restricted model, the random out-regular simple graphs ($\mathrm{RSG}^+(s)$), with less freedom and independence of the edges. Simple graphs are more natural in real-world applications like citation graphs and k-nearest neighbors graphs where self-loops and parallel edges are not allowed. We prove random walks on a $\mathrm{RSG}^+(s)$ converge to the stationary distribution polynomially fast with probability $1 - o(1)$. Based on the proofs for out-regular models, we then show similar properties for in-regular models. In-regular graphs are less popular and of limited interest in practice but their properties are helpful to studying the random s-in s-out graph models, first introduced by Fenner and Frieze [12], which can be viewed as the sum of a random out-regular graph and a random in-regular graph. After that we study the two classes of regular graphs in usual sense: regular digraphs and regular undirected graphs. They are the most restricted graphs among these models but very widely studied in the literature. Undirected sparse ($s = O(1)$) regular multigraphs are known as expander graphs. It is well known that expander graphs have well-bounded Laplacian eigenvalues. In this paper $\mathrm{RDG}(s)$ and $\mathrm{RG}(s)$ are simple, not necessarily sparse graphs. In addition, polynomially bounding the Laplacian eigenvalues for $\mathrm{RDG}(s)$ and $\mathrm{RG}(s)$ is not hard and doesn't involve any randomness (including nonsparse cases, see Appendix C.2), but bounding Laplacian eigenvalues is not sufficient for fast convergence. Most of our effort is spent on the aperiodicity, where the randomness in the models is formally dealt with. To the best of our knowledge, no work has been done on the ergodicity and convergence rate of the random walks on $\mathrm{RDG}(s)$ and previous results for $\mathrm{RG}(s)$ require $s = \lfloor \log^C n \rfloor$ for some constant $C \geq 2$, where n is the size of the graph. We present a complete proof for fast convergence of random walks on $\mathrm{RDG}(s)$ for $s \geq 2$ and random walks on $\mathrm{RG}(s)$ for $s \geq 3$ if n is odd and for $3 \leq s = o(\sqrt{n})$ or $s > \frac{1}{2}n$ if n is even.

Angluin and Chen [1] proposed a random-walk based algorithm for learning random DFAs. Observing the connection between DFA learning and label-guided graph exploration, along with the fast convergence results we prove in this paper, we generalize Angluin and Chen's algorithm to learning random regular graphs of fixed out-degree s. The learning model we use is Kearns' statistical query model [17], a weaker variant of Valiant's PAC learning model.

Random regular graph model	**Notation**
Random out-regular multigraph	$\mathrm{RMG}^+(s)$
Random out-regular simple graph	$\mathrm{RSG}^+(s)$
Random in-regular multigraph	$\mathrm{RMG}^-(s)$
Random in-regular simple graph	$\mathrm{RSG}^-(s)$
Random s-in s-out multigraph	$\mathrm{RMG}^\pm(s)$
Random s-in s-out simple graph	$\mathrm{RSG}^\pm(s)$
Random regular digraph	$\mathrm{RDG}(s)$
Random regular undirected graph	$\mathrm{RG}(s)$

Table 1: Random regular graph models with fixed degree s

Learning graphs from exploration is a long studied theoretical learning problem [3, 4], where the graphs are usually assumed out-regular. We follow [4]s settings but in the passive learning scenario where blind agents passively explore the graph on random paths. In a regular graph of out-degree s, the s edges incident from a node are associated to s distinct port numbers in $\{1, 2, \ldots, s\}$ in a one-to-one manner, which is a standard label-guided graph exploration setting ([4, 14, 22]). Each edge of a node is labeled with its local port number. The input data to the statistical query oracle are of the form (x, v) where x is a random uniform path (a sequence of edge labels) of a fixed length and v is the vertex of the graph reached on the path x starting from a particular start vertex v_0.

Related works

The study of random regular graphs started with the works of Bender [2], Bollobás [5] and Wormald [28]. Their applications in computer science soon led to a large volume of subsequent works in this area (see [29] for a survey). Most of these contributions are on the topics of asymptotic enumeration, chromatic number and Hamilton cycles. Nevertheless, research on random walks on random regular graphs is very limited in the literature. Hildebrand [16] showed the fast convergence of random walks on a $\mathrm{RG}(s)$ with the constraint $s = \Theta(\log^C n)$ for some constant $C > 2$ and Cooper [10] studied the cover time with fixed constant $s = O(1)$ but no convergence result was presented. In the context of DFA learning, Angluin and Chen [1] first proved the fast convergence of random walks on a $\mathrm{RMG}^+(s)$ for $s \geq 2$.

The first known algorithm designed for graph exploration was introduced by Shannon [25]. Since then, many subsequent works have studied the feasibility of graph exploration in the port numbering setting. Rollik [23] gave a complete proof of that no robot with a finite number of pebbles can explore all graphs. The result holds even when restricted to planar 3-regular graphs. Without pebbles, it was proved [14] that a robot needs $\Theta(Diam \cdot \log s_0)$ bits of memory to explore all graphs of diameter $Diam$ and maximum degree s_0.

2 Random walks on random regular graphs

In this section, we describe our main theoretical result. Concepts and notation used throughout this paper are described in Section 2.1. The main theorem is presented in Section 2.2 with a sketch of the proof. The details of the complete proof are deferred to the appendices.

2.1 Preliminaries

A *graph* is a tuple $G = (V, E)$, where V is a (finite) set whose elements are called *vertices* and E is a (finite) multiset of ordered pairs of V called *edges*. We denote by $n = |V|$. A graph is *undirected* if the vertex pairs in E are unordered, and is *simple* if it has no self-loops or parallel edges. If vertex v is reachable from another vertex u, the *distance* $d(u, v)$ from u to v is the minimum length of the paths from u to v and $d(u, u) = 0$. The *diameter* of a graph is $\max\{d(u, v) \mid v = u \text{ or } v \text{ is reachable from } u\}$. A graph G is *(cyclically) h-partite* if V can be partitioned into h subsets, $V_0, V_1, \ldots, V_{h-1}$, in such a way that all edges from V_i go to $V_{(i+1) \bmod h}$. We say a vertex set $V_0 \subseteq V$ is *closed* if for any $u \in V_0$ and any v such that $(u, v) \in E$, we must have $v \in V_0$. A component $V_0 \subset V$ is *isolated* if for any $u \in V_0$ and any v such that $(u, v) \in E$ or $(v, u) \in E$, we must have $v \in V_0$.

In an undirected graph, the *degree* of a vertex u is the number of edges incident to u. An undirected graph is *regular* if every vertex has the same degree. In a digraph, for a directed edge (u, v) in E, we say that vertex u has an *out-neighbor* v and vertex v has an *in-neighbor* u. The number of edges incident to a vertex u is the *in-degree* of u, denoted by d^-_u, and the number of edges incident from u is its *out-degree*, denoted by d^+_u. Unless otherwise stated, by default a *neighbor* refers to an out-neighbor and the degree of a vertex u denoted by d_u means the out-degree. A graph G is *out-regular* if $d_u = s$ for $\forall u \in V$; and is *in-regular* if $d^-_u = s$ for $\forall u \in V$. A digraph is *regular* if it is both in-regular and out-regular.

A *walk* on a graph G is a sequence of vertices $(v_0, v_1, \ldots, v_\ell)$ such that $(v_{i-1}, v_i) \in E$ for all $1 \leq i \leq \ell$. A *random walk* on a graph G is defined by a transition probability matrix P with $P(u, v) = \#\{(u, v) \in E\} \cdot d_u^{-1}$ denoting the probability of moving from vertex u to vertex v, where $\#\{(u, v) \in E\}$ is the number of edges from u to v in the graph. A vertex (or equivalently a state of a random walk) u is *aperiodic* if $\gcd\{t \geq 1 \mid P^t(u, u) > 0\} = 1$. A graph G (or a random walk on G) is *irreducible* if for every u and v in V there exist a directed cycle in G containing u and v, and is aperiodic if every vertex is aperiodic. A distribution vector ϕ satisfying $\phi P = \phi$ is called a *Perron* *vector* of the walk. An irreducible and aperiodic random walk has a unique Perron vector ϕ and $\lim_{t\to+\infty} P^t(u,\cdot) = \phi$ (called the *stationary distribution*) for any $u \in V$. In the study of rapidly mixing walks, the *convergence rate* in the L_2 distance $\Delta_{L_2}(t) = \max_{u\in V} \|P^t(u,\cdot) - \phi\|_2$ is often used. A stronger notion in L_1 distance is measured by the *total variation distance*, given by $\Delta_{TV}(t) = \frac{1}{2}\max_{u\in V}\sum_{v\in V}|P^t(u,v) - \phi(v)|$. Another notion of distance for measuring convergence rate is the χ*-square distance*:

$$\Delta_{\chi^2}(t) = \max_{u\in V}\left(\sum_{v\in V}\frac{\left(P^t(u,v) - \phi(v)\right)^2}{\phi(v)}\right)^{\frac{1}{2}}$$

As the Cauchy-Schwarz inequality gives $\Delta_{L_2}(t) \le 2\Delta_{TV}(t) \le \Delta_{\chi^2}(t)$, a convergence upper bound for $\Delta_{\chi^2}(t)$ also bounds $\Delta_{L_2}(t)$ and $\Delta_{TV}(t)$.

In this paper we study the random graph models listed in Table 1. For each model, an instance is drawn uniformly at random from the instance space of the model. A random s-in s-out graph is generated as the sum of a random in-regular graph and a random out-regular graph [12]. A RDG(s) has no parallel edges but allows self-loops. A RG(s) is simple.

2.2 The main theorem

We prove positive results on the ergodicity and convergence rate of random walks on random regular graphs, as stated in the following theorem.

Theorem 1 *With probability* $1 - o(1)$*, a random walk on a random regular graph has* $\Delta_{\chi^2}(t) \le e^{-k}$ *after* $t \ge t_0$ *steps, where*

1. *for* $RMG^+(s)$ *and* $RSG^+(s)$*:* $t_0 = 2C(C+1)sn^{1+C}(\log n + k)\cdot \log_s n$ *for some constant* $C > 0$ *when* $s \ge 2$*;*
2. *for RDG(s):* $t_0 = 2s(n-1)(\log n + 2k)$ *when* $s \ge 2$*;*
3. *for RG(s):* $t_0 = s(n-1)(\log n + 2k)$ *when* $s \ge 3$ *if* n *is odd; when* $3 \le s = o(\sqrt{n})$ *or* $s > \frac{1}{2}n$ *if* n *is even;*
4. *for* $RMG^-(s)$ *and* $RSG^-(s)$*:* $t_0 = 2C(C+1)s^C n^{1+C}(\log n + k)\cdot\log_s n$ *for some constant* $C > 0$ *when the walk is restricted to the unique irreducible component and there exists a constant* $C' \ge 1$ *such that* $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$.
5. *for* $RMG^\pm(s)$ *and* $RSG^\pm(s)$*:* $t_0 = 2C(C+1)s^C n^{1+C}(\log n + k)\cdot\log_s n$ *for some constant* $C > 0$ *when there exists a constant* $C' \ge 1$ *such that* $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$.

The constraints on s in the theorem are reasonable. The low connectivity of 1-regular graphs makes them of little interest so we need at least $s \geq 2$. In the undirected case we have $s \geq 3$ because when $s = 2$ a connected 2-regular undirected graph (or component) can only be a simple cycle. That is, a RG(2) must be a set of isolated simple cycle(s). This not only breaks the irreducibility, but also violates the aperiodicity. The other constraint $s = o(\sqrt{n})$ for even n comes from the study of enumeration on RG(s). In the cases (4) and (5), a lower bound on s is needed because small in-degree s brings us large maximum out-degree (with respect to s). Unlike other models, the irreducible component in the in-regular cases in the theorem is not necessarily closed, and the fast convergence property only holds when the walk is restricted to the unique irreducible component. We present the main idea of the proof with most details deferred to the appendices.

2.3 Fast convergence on RMG$^+(s)$ and RSG$^+(s)$

Angluin and Chen [1] first proved that random walks on a random DFA converge polynomially fast. Because the underlying graph of a random DFA exactly is a RMG$^+(s)$, the RMG$^+(s)$ case in the main theorem is established immediately by their work.

In Appendix A we present a complete proof for the RSG$^+(s)$ case. A standard proof of fast convergence consists of three parts: irreducibility, aperiodicity and polynomial convergence rate. The irreducibility of RSG$^+(s)$ is built on that of RMG$^+(s)$, thanks to the similarities they share. A RSG$^+(s)$ can be generated from a RMG$^+(s)$ using a two-stage procedure. Stage 1: generate a RMG$^+(s)$. Stage 2: for each vertex in the graph, check whether all its s neighbors are distinct nodes that are not itself. If not, keep choosing neighbors from V uniformly at random until it has exactly s distinct neighbors excluding itself. Finally, remove self-loops and merge parallel edges to simple edges. Since a RMG$^+(s)$ can be viewed as a RMG$^+(s)$ adding more edges after removing self-loops and merging parallel edges, together with the fact that a RMG$^+(s)$ has a large closed and strongly connected component (Lemma 2), we achieve the irreducibility of RSG$^+(s)$ (Lemma 3).

Let $p_h(\bar{n})$ be the probability of existence of an h-partite component of size $\bar{n}$, denoted by $\bar{G} = (\bar{V}, \bar{E})$, in a RSG$^+(s)$. $\bar{G}$ is h-partite if and only if $\bar{V}$ can be partitioned into h subsets, $\bar{V}_0, \bar{V}_1, \ldots, \bar{V}_{h-1}$, such that all edges from $\bar{V}_i$ go to $\bar{V}_{(i+1) \mod h}$. Algebra and combinatorics bounds give us

$$p_h(\bar{n}) \leq \binom{n}{\bar{n}} \cdot \left(\frac{1}{h}\right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1}\right)^{2\bar{n}}$$

and that $p_h(\bar{n})$ is exponentially small for any $\bar{n} > 0.79n$ and $h \geq 2$ so that the probability of periodicity $\leq \sum_{\bar{n}=\lceil 0.79n \rceil}^{n} \sum_{h=2}^{\bar{n}} p_h(\bar{n})$ goes to 0 when $n \to +\infty$ (Lemma 4).

The proof of the polynomial convergence rate is mainly done by showing that a RSG$^+(s)$ has logarithmic diameter (of order $\Theta(\log_s n)$) with high probability. Similar to the proof of the same argument for RMG$^+(s)$ [27], we generate a RSG$^+(s)$ in a "level-wise" order. Pick a vertex $u_0 \in V$ and let level 0 be $\{u_0\}$. Inductively, for each vertex in level $i-1$ we choose its s neighbors uniformly excluding itself without replacement. All the new chosen vertices form level i. We call the set of vertices in level $\leq i$ the ball i. To accomplish the proof, we divide the above spanning procedure into six stages (see the proof of Theorem 4 for details). We show that the size of the spanning ball keeps increasing in the first 3 stages while the boundary of the ball starts shrinking in Stage 4 and finally the spanning procedure halts with an empty new level. The number of levels constructed in every stage is logarithmic, and so is the diameter of the graph.

2.4 Fast convergence on RMG$^-(s)$, RSG$^-(s)$, RMG$^\pm(s)$ and RSG$^\pm(s)$

The conclusions drawn for random walks on random out-regular graphs can be easily generalized to the in-regular cases. Let A be the adjacency matrix of graph G. Denote by $G^\top$ the *transpose* of G defined by adjacent matrix $A^\top$. It is apparent to see that (1) $G^\top$ has exactly the same irreducible components as G; (2) The aperiodicity of G implies the aperiodicity of $G^\top$; (3) $Diam(G) = Diam(G^\top)$. These give irreducibility, aperiodicity and logarithmic diameter. Note that in these cases the irreducible component is usually not closed. Hence, the fast convergence argument only holds when the walk is restricted to the unique irreducible component. According to Theorem 3 (in Appendix A), it remains to bound the maximum out-degree $s_0 = \arg\max_{u\in V} d_u$. This requires the lower bound assumption on s as stated in the main theorem because small in-degree s results in large maximum out-degree (with respect to s). A random s-in s-out graph can be viewed as the sum of a random out-regular graph and a random in-regular graph, generated independently of each other. Thus logarithmic diameter is trivial. The original paper by Fenner and Frieze [12] has already shown the strong connectivity of the random s-in s-out graphs for $s \geq 2$. As the entire graph is strongly connected, the connected component is surely closed and unique. Their aperiodicity is established by the fact that sum graph retains all directed cycles in the original graphs. Please refer to Appendix B for the complete proof.

2.5 Fast convergence on RDG(s) and RG(s)

Among all the models in this paper, RDG(s) is the most constrained one, due to the strong dependence and strict constraints on the edge selections (same in the undirected model) and the lack of symmetry (while the undirected model has symmetry). Unlike the previous cases, the proof is based on enumeration.

Previous works have contributed the irreducibility. The proof of aperiodicity starts with the asymptotic enumeration of regular digraphs. Note the bijection between regular digraphs and binary square matrices with equal line sums. Let $N(n, s)$ be the number of s-regular digraphs of size n. We present an asymptotic formula for $N(n, s)$, by unifying previous results on binary square matrices with equal line sums (Lemma 8). We also observe the bijection between regular digraphs of size n and colored regular bipartite (undirected) graphs of size $2n$. Let $G = (V, E)$ be a regular digraph of fixed degree s. We construct a regular bipartite graph $G' = (V', E')$ where $|V'| = 2|V|$ as following. Without loss of generality, denote $V = \{v_1, v_2, \ldots, v_n\}$ and $V' = \{v'_1, v'_2, \ldots, v'_{2n}\}$ with $\{v'_1, v'_2, \ldots, v'_n\}$ of one color and $\{v'_{n+1}, v'_{n+2}, \ldots, v'_{2n}\}$ of the other. Let $(v'_i, v'_{n+j}) \in E'$ if and only if $(v_i, v_j) \in E$. We can see such a regular bipartite graph G' is unique for each regular digraph G and vice versa.

To show aperiodicity, we again need to exponentially upper-bound the probability of the graph being h-partite, denoted by p_h. If V can be partitioned into h subsets, $V_0, V_1, \ldots, V_{h-1}$, such that all edges from V_i go to $V_{(i+1) \mod h}$, because the graph is regular, we must have $|V_0| = |V_1| = \ldots = |V_{h-1}| = \frac{n}{h}$ and $h \leq \frac{n}{s}$. Notice that the number of possible edge combinations from V_i going to $V_{(i+1) \mod h}$ is exactly the number of colored s-regular bipartite (undirected) graphs of size $\frac{n}{h}$, which is $N(\frac{n}{h}, s)$. This gives

$$p_h \leq \frac{1}{h} \cdot \binom{n}{\frac{n}{h}\ \frac{n}{h} \cdots \frac{n}{h}} \cdot \frac{\left[N\left(\frac{n}{h}, s\right)\right]^h}{N(n, s)}$$

With the asymptotic enumeration result we complete the proof (Lemma 9).

Unlike the previous cases where we prove fast convergence by proving logarithmic diameter, for regular digraphs the polynomial convergence rate follows from a lower bound on the first non-zero eigenvalue on the Laplacian matrix. Note that the walk matrix $P = \frac{1}{s}A$ of a random walk on a RDG(s) is doubly stochastic matrix, and so is $\frac{1}{2}(P + P^\top)$. Also observe that the Perron vector of any regular digraph is always the uniform distribution. Using a spectral lower bound for doubly stochastic matrices due to Fiedler [13], we complete the proof.

Random regular undirected graphs are much more widely studied than directed ones, mainly owing to the symmetry of undirected graphs. Previous works have established connectivity and enumeration results. Because the only periodic case for an undirected graph is being bipartite, we only need to bound the probability p_2. This is again done by enumeration. From the proof in the preceding case we already know the number of bipartite s-regular undirected graphs of size n is $\binom{n}{\frac{n}{2}} \cdot N(\frac{n}{2}, s)$. Denote by $N'(n, s)$ the number of s-regular undirected graphs of size n. We have

$$p_2 \leq \frac{1}{2}\binom{n}{\frac{n}{2}} \cdot \frac{N(\frac{n}{2}, s)}{N'(n, s)}$$

Using the same spectral lower bound for doubly stochastic matrices as in the preceding case, we have the polynomial convergence rate. Detailed algebra is deferred to Appendices C and D.

3 Reconstructing random regular graphs from random paths

The positive theoretical result in Section 2 establishes the generalization of Angluin and Chen's algorithm to learning random regular graphs. Because the nature of the algorithm requires the graph to be out-regular, we only apply this algorithm to the models with fixed out-degree s, namely RMG$^+(s)$, RSG$^+(s)$, RDG(s) and RG(s).

3.1 Preliminaries

In a computational learning model, an algorithm is usually given access to an oracle providing information about the target concept. Kearns [17] modified Valiant's model and introduced the *statistical query oracle STAT*. Kearns' oracle takes as input a statistical query of the form (χ, τ). Here χ is any mapping of a labeled example to $\{0, 1\}$ and $\tau \in [0, 1]$ is called the noise *tolerance*. Let c be the target concept and $\mathcal{D}$ be the distribution over the instance space. Oracle $STAT(c, \mathcal{D})$ returns to the learner an estimate for the expectation $\mathbf{E}\chi$, that is, the probability that $\chi = 1$ when the labeled example is drawn according to $\mathcal{D}$. A statistical query can have a condition, in which case $\mathbf{E}\chi$ is a conditional probability. This estimate is accurate within additive error τ. Kearns [17] proved that the statistical query model is weaker than the classic PAC model. That is, PAC learnability from oracle *STAT* implies PAC learnability from the classic example oracle, but not vice versa.

In this section we study the problem of learning regular graphs in the statistical query model. In a typical label-guided graph exploration setting ([3, 4, 14, 22]), in a regular graph with fixed out-degree s, the s edges incident from a node are associated to s distinct *port numbers* in $\Sigma = \{1, 2, \ldots, s\}$, in a one-to-one manner. Each edge of a node is labeled with the associated port number. Port numbering is *local*, i.e., there is no relation between port numbers at u and at v. In the undirected case RG(s), every undirected edge (u, v) has two labels corresponding to its port numbers at u and at v respectively. A *path* is a sequence of edge labels. The input data to the statistical query oracle are of the form (x, v) where $x \in \Sigma^t$ is a random uniform path and v is the vertex of the graph reached on the path x starting from a particular *start vertex* v_0. Here $t = poly(n, s)$ is the length of the example paths. The learner has access to the oracle *STAT* and algorithms are designed to reconstruct the graph (or the unique closed irreducible component for RMG$^+(s)$ and RSG$^+(s)$).

3.2 The learning algorithm

A uniform path $x \in \Sigma^t$ corresponds to a random walk of length t on the graph G starting from the start vertex v_0. Since all these four types of random regular graphs have been proved to have one unique closed irreducible component with high probability and due to the main theorem, the walk will converge to the stationary distribution p_λ polynomially fast, with any start vertex. Define a collection of $n \times n$ binary matrices M_σ indexed by labels $\sigma \in \Sigma$ as follows. For each pair of vertices (u, v), the element $M_\sigma(u, v)$ is 1 if $(u, v) \in E$ and is labeled with σ at vertex u, and 0 otherwise. For a path $y = y_1 y_2 \ldots y_m$ of length m, define M_y to be the matrix product $M_y = M_{y_1} \cdot M_{y_2} \ldots M_{y_m}$. Also define the distribution vector p_y over V obtained by starting with the stationary distribution p_λ and walking along the path y on the graph. That is, $p_y = p_\lambda M_y$. Note that here we use notation y to distinguish a general path of length m from the example paths x of length t. Let z be the i-th column of matrix M_σ, P_A be the $s^{\Theta(\log_s n)} \times n$ coefficient matrix whose rows are $\{p_y \mid y \in \Sigma^{\Theta(\log_s n)}\}$ and b be the vector consisting of $\{p_{y\sigma}(i) \mid y \in \Sigma^{\Theta(\log_s n)}\}$ corresponding to each y in P_A. The algorithm recovers the strongly connected component by solving the linear equation system $P_A z = b$ for each column z in each matrix M_σ.

By setting $k = \log \frac{2}{\tau}$ in the main theorem, after t_0 steps the random walk converges to the stationary distribution p_λ within χ-square distance $\frac{\tau}{2}$ with high probability. Observe that $2\|\phi_t - \phi\|_{TV} \le \Delta_{\chi^2}(t)$, where ϕ_t is the distribution vector over V after t steps of random walk. We can estimate the stationary distribution for a vertex i by the fraction of examples (x, v) such that $v = i$. In general, for any path y, we can estimate the value of p_y for a vertex i as the ratio between the number of pairs (x, v) such that y is a suffix of x and $v = i$ and the number of examples (x, v) where y is a suffix of x. In the statistical query model this is done with a conditional statistical query $\chi_{y,i}(x, v) = \mathbb{1}\{v = i \mid y \text{ is a suffix of } x\}$ at tolerance $\frac{\tau}{2}$, where $\mathbb{1}$ is the boolean indicator function. Denote by vector $\widehat{p}_y$ the query result returned by oracle $STAT$ where $\widehat{p}_y(i)$ is the estimate $\mathbf{E}\chi_{y,i}$, and by $\widehat{P}_A$ and $\widehat{b}$ the estimates for P_A and b respectively from oracle $STAT$. We have $\|p_y - \widehat{p}_y\|_\infty \le \tau$ for any path y [1]. The algorithm approximates z by solving the perturbed linear least squares problem: $\min_z \|\widehat{P}_A z - \widehat{b}\|_2$. Let vector $\widehat{z}$ be the solution. Then from [1] we have

Lemma 1 *If P_A has full rank with high probability, for all columns z in all matrices M_σ, $\|z - \widehat{z}\|_\infty \le \|z\|_1 \||P_A^\dagger|\|_\infty \tau + O(\tau^2)$ with probability $1 - o(1)$.*

For RMG$^+(s)$, it is proved in [1] with high probability $\|z\|_1 \le \frac{(1+\varepsilon)\log ns}{\log\log ns}$ for any constant $\varepsilon > 0$. We show this also holds for RSG$^+(s)$ (see Appendix E). For RDG(s) and RG(s), we have $\|z\|_1 = s$.

Theorem 2 *If P_A has full rank with high probability,*

1. *for RMG$^+(s)$ and RSG$^+(s)$, $\|z - \widehat{z}\|_\infty \leq \frac{(1+\varepsilon)\log ns}{\log\log ns}|||P_A^\dagger|||_\infty \tau + O(\tau^2)$ for any constant $\varepsilon > 0$*
2. *for RDG(s) and RG(s), $\|z - \widehat{z}\|_\infty \leq s|||P_A^\dagger|||_\infty \tau + O(\tau^2)$*

holds for all columns z in all matrices M_σ with probability $1 - o(1)$.

This further implies that if we set the tolerance $\tau = \frac{\log\log ns}{3|||P_A^\dagger|||_\infty \log ns}$ for RMG$^+(s)$ and RSG$^+(s)$, and $\tau = \frac{1}{3s|||P_A^\dagger|||_\infty}$ for RDG(s) and RG(s), the solution error $\|z - \widehat{z}\|_\infty < \frac{1}{2}$ with high probability. Based on the prior knowledge we have on z, we could refine $\widehat{z}$ by rounding up $\widehat{z}$ to a binary vector $\tilde{z}$, i.e., for each $1 \leq i \leq n$, $\tilde{z}(i) = 1$ if $\widehat{z}(i) > \frac{1}{2}$ and 0 otherwise, whereby we will have $\tilde{z}(v) = z(v)$ for any vertex v. We provide a toy example in Appendix G to demonstrate how the learning algorithm works on a concrete regular graph.

3.3 Experiments and empirical results

In this section we present experimental results to study the empirical performance of the learning algorithm, which was run in MATLAB on a workstation built with Intel i5-2500 3.30GHz CPU and 8GB memory. To be more robust against fluctuation from randomness, each test was run for 20 times and the medians were taken. The graphs are generated uniformly at random as defined and the algorithm solves the equation system $\{p_y M_\sigma = p_{y\sigma} \mid y \in \Sigma^{\leq \lceil \log_s n \rceil}\}$ using the built-in linear least squares function in MATLAB. We simulate the statistical query oracle with uniform additive noise from $[-\tau, \tau]$. Since Angluin and Chen's paper [1] already included experiments on learning a random DFA, whose underlying graph is exactly RMG$^+(s)$, we don't duplicate the experiments for RMG$^+(s)$. As this is a theoretical paper, we defer all detailed experimental results to Appendix H.

The generating procedure of a RSG$^+(s)$ is standard. Each node $v \in V$ independently chooses s neighbors without replacement uniformly at random. However, to the best of our knowledge, there is no algorithm that efficiently generates a RDG(s) or a RG(s). In our experiments, we use the celebrated pairing model first introduced by Bollobás [5]. In a RG(s), each vertex has s *ports* associated to its s edges. It is well known that the necessary and sufficient conditions for an s-regular graph of order n to exist are that $n \geq s + 1$ and that ns is even. To generate a RG(s), we uniformly pick a perfect matching of the ns ports into $\frac{1}{2}ns$ pairs. Adding an edge between each pair of ports gives a (not necessarily simple) regular graph. Repeat this procedure until it produces a simple graph. Likewise we generate a RDG(s) by uniformly matching ns out-ports (corresponding to outgoing edges) with ns in-ports (corresponding to incoming edges) until we get a regular digraph with no parallel edge. This method is not efficient owing to the unbounded number of repetitions, especially when s grows. Hence, with large s this generating method is extremely slow.

Note that this limitation comes from the existing generating methods. Our learning algorithm is efficient.

The experiments start with an empirical estimate for the norm $|||P_A^\dagger|||_\infty$. For $\mathrm{RSG}^+(s)$ we first vary the graph size n from 32 to 4300 with fixed out-degree $s = 2$. Figure 2 shows the curve of $|||P_A^\dagger|||_\infty$ versus n with fixed s. Notice that the threshold phenomenon in the plot comes from the ceiling operation in the algorithm configuration. When n is much smaller than the threshold $s^{\lceil \log_s n \rceil}$, the system is overdetermined with many extra equations. Thus it is robust to perturbation and well-conditioned. When n approaches the threshold $s^{\lceil \log_s n \rceil}$, the system has fewer extra equations and becomes relatively more sensitive to perturbations, for which the condition number increases until the graph size reaches $n = s^i$ for the next integer i. One can avoid this threshold phenomenon by making the size of the equation system grow smoothly as n increases. We then fix n to be 256 and vary s from 2 to 75, as shown in Figure 3. Similarly there is the threshold phenomenon resulting from the ceiling strategy. All peaks where $n = s^i$ are included and plotted. Meanwhile the rank of the coefficient matrix P_A is measured to support the full-rank assumption. Both figures suggest an upper bound $ns \log s$ for $|||P_A^\dagger|||_\infty$ of $\mathrm{RSG}^+(s)$. Figures 8 and 9 demonstrate the experimental results for the maximum absolute error. Along with the error curve a function is plotted to approximate the order of the decline rate of the error. An empirical error bound is $O(\log^{-1} n)$ with fixed s and $O(1/\sqrt{s})$ with fixed n.

Because generating a $\mathrm{RDG}(s)$ and generating a $\mathrm{RG}(s)$ are extremely slow with large s, the range of s where we can efficiently conduct the experiments is very limited. For $\mathrm{RDG}(s)$ we first vary n from 32 to 4300 with fixed $s = 2$ (Figure 4) as before but with fixed $n = 256$ we vary s from 2 to 6 (Figure 5). The norm $|||P_A^\dagger|||_\infty$ of $\mathrm{RDG}(s)$ is bounded by $n \log^3(ns)$ and an empirical error bound is $O(\log^{-1} n)$ with fixed s (Figure 10) and $O(1/s)$ with fixed n (Figure 11). For $\mathrm{RG}(s)$ we vary n from 26 to 3000 with fixed $s = 3$ (Figure 6) and vary s from 3 to 8 with fixed $n = 242$ (Figure 7). As the existence of a regular undirected graph requires even ns and s is fixed to be 3 when varying n, we only run experiments with even n. For critical points where $n = 3^i$, experiments are run with $n = 3^i - 1$ and $n = 3^i + 1$. This explains why we start with $n = 26$ instead of $n = 27$ with fixed $s = 3$, and also why we fix $n = 242$ rather than $n = 243$ when varying s. The norm $|||P_A^\dagger|||_\infty$ of $\mathrm{RG}(s)$ is bounded by $sn^{1.6}$ and an empirical error bound is $O(\log n/\sqrt{n})$ with fixed s (Figure 12) and $O(1/s)$ with fixed n (Figure 13).

4 Other applications and discussion

With the broad applications of regular graphs in computer science and machine learning, our theoretical results can be applied to other research areas such as distributed networks and social network graphs. Performing random walks on distributed networks is an active area of research (see [7] for a survey). High connectivity, bounded degree and low diameter are very common properties of (well designed) distribution network models. Theorem 3 explicitly provides fast convergence for random walks on these models. For instance, Pandurangan et al. [21] proposed a protocol which ensures that the network is connected and has logarithmic diameter with high probability, and has always bounded degree. A simpler, fully decentralized model named SWAN was proposed by Bourassa and Holt [6] based on random walks, which produces a random regular graph. In another direction, random walks have proven to be a simple, yet powerful mathematical tool for extracting information from large scale and complex social networks (see [24] for a survey). Social network graphs also have the above properties (high connectivity, small degree and low diameter) so that the random walks shall converge fast as we proved. One application of fast convergence is the capability of uniformly sampling the graph, which is very important in many graph learning problems.

In this paper we have shown positive theoretical results on random walks on random regular graphs, and generalized Angluin and Chen's algorithm to learning random regular graphs from random paths. One technical question on the fast convergence result is whether it can be generalized to weighted random walks on random regular graphs. An immediate benefit from this generalization is the release from the requirement of uniform paths in the learning algorithm. However, we conjecture this requires a polynomial lower bound on the edge weights in the graph, to avoid exponentially small nonzero elements in the walk matrix P. Another potential future work is to apply this algorithm to learning a more general class of graphs. Note that any generalization of the algorithm needs not only fast convergence, but also asymmetry of the target graph. The class of permutation automata [26] is one example that has symmetric graph structure and degenerate P_A. Also, there is potential possibility of relaxing the constraint on s in the RG(s) case if advances on the enumeration of regular undirected graphs are made.

Appendix A Proof of Theorem 1 for RSG$^+(s)$

In this section we prove the fast convergence of random walks on RSG$^+(s)$, divided into three parts: irreducibility, aperiodicity and polynomial convergence rate.

A.1 Irreducibility

Since RMG$^+(s)$ and RSG$^+(s)$ share many similarities, we can achieve the irreducibility of RSG$^+(s)$ based on that of RMG$^+(s)$.

Lemma 2 ([15]) *With probability $1 - o(1)$, a RMG$^+(s)$ has a unique strongly connected component, denote by $\tilde{G} = (\tilde{V}, \tilde{E})$, of size $\tilde{n}$, and a)* $\lim_{n\to+\infty} \frac{\tilde{n}}{n} = C$

for some constant $C > 0.7968$ when $s \geq 2$ or some $C > 0.999$ when $s \geq 7$; b) $\tilde{V}$ is closed.

The irreducibility of RSG$^+(s)$ is proved in the following lemma.

Lemma 3 *With probability $1-o(1)$, a RSG$^+(s)$ has a unique closed and strongly connected component, denoted by $\tilde{G} = (\tilde{V}, \tilde{E})$, of size $\tilde{n}$ when $n \to +\infty$, and $\lim_{n\to+\infty} \frac{\tilde{n}}{n} \geq C$ for some constant $C > 0.7968$ when $s \geq 2$ or some $C > 0.999$ when $s \geq 7$.*

Proof Recall that the only difference of RSG$^+(s)$ from RMG$^+(s)$ is that the s neighbors of each vertex are chosen without replacement so no self-loops or parallel edges are allowed. We can consider the following two-stage procedure to generate a RSG$^+(s)$ from a RMG$^+(s)$. Stage 1: generate a RMG$^+(s)$. Stage 2: for each vertex in the graph, check whether all its s neighbors are distinct nodes that are not itself. If not, keep choosing neighbors from V uniformly at random until it has exactly s distinct neighbors excluding itself. Finally, remove self-loops and merge parallel edges to simple edges. Because each $v \in V \setminus \{u\}$ will become u's neighbor with equal probability, the result graph of this procedure is a uniformly generated RSG$^+(s)$.

Thus a RMG$^+(s)$ can be viewed as a RMG$^+(s)$ adding more edges after removing self-loops and merging parallel edges. This means the simple graph model has better connectivity. The size of the strongly connected component will only increase. After Stage 1 we have a RMG$^+(s)$, denoted by $G_1 = (V, E_1)$ and let $\tilde{V_1} \subseteq V$ be the closed strongly component of G_1 stated in Lemma 2. To show the irreducible component in a RSG$^+(s)$ is also closed, note that for any $v \notin \tilde{V_1}$, there must exist at least one path from v to $\tilde{V_1}$. Otherwise there will be another strongly connected component in G_1, which contradicts Lemma 2. Thus in Stage 2, every time when we add an edge from $\tilde{V_1}$ to some $u \notin \tilde{V_1}$, there must be some directed path(s) from u heading back to the irreducible component. All the vertices on this(these) path(s) are now strongly connected with $\tilde{V_1}$ and become new members of the irreducible component. Therefore, the irreducible component in the final simple graph will also be closed. ■

A.2 Aperiodicity

Lemma 4 *With probability $1 - o(1)$, $\tilde{G}$ in Lemma 3 is aperiodic.*

Proof Let $p_h(\bar{n})$ be the probability of existence of an h-partite component of size $\bar{n}$ in a RSG$^+(s)$. The proof is completed by showing $p_h(\bar{n})$ goes to 0 exponentially fast when $n \to +\infty$ for any $\bar{n} > 0.79n$ and $h \geq 2$ so that combining with Lemma 3 the probability of periodicity is $\leq \sum_{\bar{n}=\lceil 0.79n \rceil}^{n} \sum_{h=2}^{\bar{n}} p_h(\bar{n})$ and goes to 0 when $n \to +\infty$.

Let $\bar{G} = (\bar{V}, \bar{E})$ be a fixed component of size $\bar{n}$ in the graph. $\bar{G}$ is h-partite if $\bar{V}$ can be partitioned into h subsets, $\bar{V}_0, \bar{V}_1, \ldots, \bar{V}_{h-1}$, such that all edges from $\bar{V}_i$ go to $\bar{V}_{(i+1) \mod h}$. The number of such partitions is at most $h^{\bar{n}}$. The probability of forming a particular partition $\bar{V}_0, \bar{V}_1, \ldots, \bar{V}_{h-1}$ is

$$\begin{aligned}
\prod_{i=0}^{h-1} \left(\frac{\binom{|\bar{V}_{(i+1) \mod h}|}{s}}{\binom{n-1}{s}} \right)^{|\bar{V}_i|} &= \prod_{i=0}^{h-1} \left(\frac{\prod_{j=0}^{s-1} \left(|\bar{V}_{(i+1) \mod h}| - j \right)}{\prod_{j=0}^{s-1} (n-1-j)} \right)^{|\bar{V}_i|} \\
&\leq \prod_{i=0}^{h-1} \left(\prod_{j=0}^{s-1} \frac{|\bar{V}_{(i+1) \mod h}|}{n-1} \right)^{|\bar{V}_i|} \\
&= \prod_{i=0}^{h-1} \left(\frac{|\bar{V}_{(i+1) \mod h}|}{n-1} \right)^{s|\bar{V}_i|} \\
&\leq \left(\frac{\bar{n}}{h(n-1)} \right)^{s\bar{n}} \\
&\leq \left(\frac{\bar{n}}{h(n-1)} \right)^{2\bar{n}}
\end{aligned}$$

This is because the product $\prod_{i=0}^{h-1} x_{(i+1) \mod h}^{x_i}$, given $x_i > 0$ and $\sum_{i=0}^{h-1} x_i = \bar{n}$, is maximized for $x_i = \bar{n}/h, i = 0 \ldots h-1$. Thus

$$\begin{aligned}
p_h(\bar{n}) \leq& \binom{n}{\bar{n}} \cdot h^{\bar{n}} \cdot \left(\frac{\bar{n}}{h(n-1)} \right)^{2\bar{n}} \\
=& \binom{n}{\bar{n}} \cdot \left(\frac{1}{h} \right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1} \right)^{2\bar{n}} \\
\leq& \binom{n}{\bar{n}} \cdot \left(\frac{1}{2} \right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1} \right)^{2\bar{n}}
\end{aligned}$$

When $\bar{n} = n$, as $\lim_{n \to +\infty} \left(\frac{n}{n-1} \right)^{2n} = e^2$, apparently $p_h(n)$ goes to 0 exponentially fast.

When $0.79n < \bar{n} < n$, we have

$$
\begin{aligned}
p_h(\bar{n}) \leq & \binom{n}{\bar{n}} \cdot \left(\frac{1}{2}\right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1}\right)^{2\bar{n}} \\
= & \frac{n!}{\bar{n}!(n-\bar{n})!} \cdot \left(\frac{1}{2}\right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1}\right)^{2\bar{n}} \\
\leq & \frac{\sqrt{2\pi n} \cdot n^n \cdot e^{n-\bar{n}+\frac{1}{12n}} \cdot e^{\bar{n}}}{\sqrt{2\pi(n-\bar{n})} \cdot e^n \cdot (n-\bar{n})^{n-\bar{n}} \cdot \sqrt{2\pi\bar{n}} \cdot \bar{n}^{\bar{n}}} \cdot \left(\frac{1}{2}\right)^{\bar{n}} \cdot \left(\frac{\bar{n}}{n-1}\right)^{2\bar{n}} \\
= & \sqrt{\frac{n}{2\pi\bar{n}(n-\bar{n})}} \cdot e^{\frac{1}{12n}} \cdot \left(\frac{\bar{n}^2}{2n^2}\right)^{\bar{n}} \cdot \frac{n^n}{\bar{n}^{\bar{n}} \cdot (n-\bar{n})^{n-\bar{n}}} \cdot \left(\frac{n}{n-1}\right)^{2\bar{n}} \\
\leq & \sqrt{\frac{n}{2\pi\bar{n}(n-\bar{n})}} \cdot e^{\frac{1}{12n}} \cdot \left(\frac{\bar{n}}{2n}\right)^{\bar{n}} \cdot \frac{n^{n-\bar{n}}}{(n-\bar{n})^{n-\bar{n}}} \cdot \left(\frac{n}{n-1}\right)^{2n} \\
= & \sqrt{\frac{n}{2\pi\bar{n}(n-\bar{n})}} \cdot e^{\frac{1}{12n}} \cdot \left[\left(\frac{\bar{n}}{2n}\right)^{\frac{\bar{n}}{n}} \cdot \left(1-\frac{\bar{n}}{n}\right)^{\frac{\bar{n}}{n}-1}\right]^n \cdot \left(\frac{n}{n-1}\right)^{2n}
\end{aligned}
$$

Note that function $f(x) = (1-x)^{x-1} \cdot \left(\frac{x}{2}\right)^x < 0.7$ for all $0.79 < x < 1$. Hence, the probability $p_h(\bar{n})$ is exponentially small, which completes the proof. ■

A.3 Fast convergence

For a random walk P on a graph, Chung [9] defined the *Laplacian matrix* $\mathcal{L}$

$$\mathcal{L} = I - \frac{\Phi^{\frac{1}{2}} P \Phi^{-\frac{1}{2}} + \Phi^{-\frac{1}{2}} P^\top \Phi^{\frac{1}{2}}}{2}$$

where Φ is an $n \times n$ diagonal matrix with entries $\Phi(u,u) = \phi(u)$. Angluin and Chen [1] proved the following theorem on convergence rate.

Theorem 3 ([1]) *A random walk on a strongly connected and aperiodic directed graph has convergence rate of order* $\Delta_{\chi^2}(t) \leq e^{-k}$ *after* $t \geq 2n \cdot Diam \cdot s_0^{1+Diam}((\log(ns_0^{Diam}) + 2k))$ *or* $t \geq 2\lambda_1^{-1}((-\log \min_u \phi(u)) + 2k)$ *steps, where Diam is the diameter of the graph,* $s_0 = \max_u d_u$ *and* λ_1 *is the smallest nonzero Laplacian eigenvalue.*

To accomplish the fast convergence of random walk on a RSG$^+(s)$, we prove the diameter of a RSG$^+(s)$ is logarithmic with high probability.

Theorem 4 *With probability* $1-o(1)$*, the diameter of a RSG*$^+$*(s) is* $\Theta(\log_s n)$*.*

Proof The logarithmic lower bound is easy to prove. For a particular vertex $u \in V$, denote by $S_i(u)$ the set of vertices in G such that for any $v \in S_i(u)$

the distance from u to v is i. We know $S_0(u) = \{u\}$ and $n = \sum_{i=0}^{+\infty} |S_i(u)|$. According to the definition of diameter, $|S_i(u)| = 0$ for all $i > Diam$. Also notice that $|S_{i+1}(u)| \leq s|S_i(u)|$, for which we have

$$n \leq 1 + s + s^2 + \ldots + s^{Diam} = \frac{s^{Diam+1} - 1}{s - 1}$$

After some algebra, $Diam \geq \log_s(n(s-1)+1) - 1 \geq \log_s(n(s-1)) - 1 = \log_s n + \log_s(s-1) - 1 \geq \log_s n - 1$ due to $\log_s(s-1) \geq 0$ for all $s \geq 2$. Hence, we have $Diam = \Omega(\log_s n)$. This lower bound holds for RMG$^+(s)$ as well.

However, the proof of the upper bound is lengthy. It is well known that a RMG$^+(s)$ has logarithmic diameter with high probability [27]. Although the proof for RMG$^+(s)$ doesn't work for RSG$^+(s)$ due to the dependence between its edge selections, our proof follows the framework of their proof.

Assume that we generate a RSG$^+(s)$ in a "level-wise" order. We pick a vertex $u_0 \in V$ and let *level* 0 be the set $\{u_0\}$. Then choose its s neighbors from $V \setminus \{u_0\}$ uniformly at random without replacement. All the neighbors of u_0 form level 1. Inductively, for each vertex in level $i-1$ we choose its s neighbors uniformly excluding itself without replacement. All the new chosen vertices form level i. We call the set of vertices in level $\leq i$ the *ball* i. By intuition, level i is the set of vertices to which the distance from u_0 is i and ball i consists of all vertices to which the distance from u_0 is at most i. Obviously level i is the boundary of ball i. The spanning procedure halts when no new vertex is chosen as a neighbor of the boundary so the next level is empty. To completely generate a RSG$^+(s)$, the final step is for each vertex not in the ball, uniformly choosing s distinct vertices as its neighbors. Let L_i be the size of level i and B_i be the size of ball i. At any time, we say a vertex is *occupied* if it has non-zero in-degree and unoccupied otherwise. During this process, *determining* a vertex refers to choosing its s neighbors.

In short, to accomplish the proof, we divide the above spanning procedure into six stages:

Stage 1 starts from the very beginning and ends at level ℓ_1 once $B_{\ell_1} \geq n^{\frac{1}{6}}$.

Stage 2 begins immediately after Stage 1 and ends at level ℓ_2 once $B_{\ell_2} \geq \frac{n}{s^4}$.

Stage 3 begins immediately after Stage 2 and ends at level ℓ_3 once $B_{\ell_3} \geq (1 - 2^{-s})\, n$.

Stage 4 begins immediately after Stage 3 and ends at level ℓ_4 once $L_{\ell_4} \leq (\log_2 n)^2$.

Stage 5 begins immediately after Stage 4 and ends at level ℓ_5 once $L_{\ell_5} \leq 120 \log_2 n$.

Stage 6 begins immediately after Stage 5 and ends at level ℓ_6 once $L_{\ell_6+1} = 0$. The spanning procedure halts.

Letting ℓ_0 be 0 and $\ell'_i = \ell_i - \ell_{i-1}$, $1 \leq i \leq 6$ be the number of new levels created in Stage i, we complete the proof by showing $\sum_{i=1}^{6} \ell'_i = O(\log_s n)$.

Now we start moving to the details. First we notice that the above level-wise procedure can also be used to generate a $\mathrm{RMG}^+(s)$ only if we choose neighbors of a vertex with replacement and allow self-loops. To distinguish between the multi-graph case and the simple graph case, let $\widehat{L}_i$ be the size of level i and $\widehat{B}_i$ be the size of ball i in the multi-graph case so that we can make use of some partial results in the multi-graph case by [27].

Consider a sequence of N Bernoulli trials with probability p for success and $1-p$ for failure. Let $X(N,p)$ denote the random variable defined as the number of successful outcomes in this sequence. [27] proved that for any $p > 0$, any natural number N and any $pN < k \leq N$, $\Pr[X(N,p) \geq k] < N \cdot [k/(pN)]^{(3+pN-k)/2}$. It's easy to see the following facts:

$$\begin{aligned}\Pr\left[X\left(ms, \frac{n-w}{n-s}\right) \leq k\right] &= \Pr\left[X\left(ms, \frac{(n-1)-(w-1)}{(n-1)-(s-1)}\right) \leq k\right] \\ &\leq \Pr[L_{i+1} \leq k \mid L_i = m \wedge B_i = w] \\ &\leq \Pr\left[X\left(ms, \frac{(n-1)-(w-1)-(ms-1)}{(n-1)-(s-1)}\right) \leq k\right] \\ &< \Pr\left[X\left(ms, \frac{n-w-ms}{n}\right) \leq k\right]\end{aligned}$$

and

$$\begin{aligned}\Pr\left[X\left(ms, \frac{n-w}{n}\right) \leq k\right] &\leq \Pr[\widehat{L}_{i+1} \leq k \mid \widehat{L}_i = m \wedge \widehat{B}_i = w] \\ &< \Pr\left[X\left(ms, \frac{n-w-ms}{n}\right) \leq k\right]\end{aligned}$$

Imagine we choose the edges one by one in the above described level-wise order. Assuming the number of occupied nodes is t at the moment when we are choosing the i-th edge of vertex v, then the probability of choosing an unoccupied vertex as the destination (so that we have a new member of the next level) is $\frac{(n-1)-(t-1)}{n-i} = \frac{n-t}{n-i}$ under the simple graph model and is always $\frac{n-t}{n}$ under the multi-graph model. Therefore, under the same configuration, we will always have higher probability to choose an unoccupied vertex under the simple graph model than that under the multi-graph model. We can easily conclude:

$$\Pr[L_{i+1} \leq k \mid L_i = m \wedge B_i = w] < \Pr[\widehat{L}_{i+1} \leq k \mid \widehat{L}_i = m \wedge \widehat{B}_i = w]$$

Similarly, imagine we determine the vertex one by one in the above described level-wise order and let $B(r)$ be the number of occupied vertices exactly after we have determined r vertices. Denote by $\widehat{B}(r)$ the corresponding quantity in the multi-graph case. From our analysis above it's easy to see $\Pr[B(r) \geq k] > \Pr[\widehat{B}(r) \geq k]$ for any $r \geq 1$. Below we will go through the six stages and show the number of new levels constructed is small in every stage.

Stage 1: For any level $i \leq \lceil \frac{1}{6} \log_s n \rceil - 1$, we have $B_{i+1} \leq \sum_{j=0}^{j+1} s^j < s^{i+2} \leq s^2 n^{\frac{1}{6}}$. Thus the probability that an edge created on level i will point to an occupied vertex is less than $\frac{s^2 n^{\frac{1}{6}}-1}{n-1} < s^2 n^{-\frac{5}{6}}$. This means that the probability that more than one edge on the first $\lceil \frac{1}{6}\log_s n\rceil - 1$ levels will point to an occupied vertex is less than $\sum_{j=2}^{k} b(j,k,p)$ where k is the maximal possible number of edges on the first $\lceil \frac{1}{6}\log_s n\rceil - 1$ levels, $p = s^2 n^{-\frac{5}{6}}$ and $b(j,k,p)$ is the probability of j successful outcomes and $k-j$ failures in k Bernoulli trials with probability p for success. Obviously, $k < s^3 n^{\frac{1}{6}}$. [27] proved that for sufficiently large n, $\sum_{j=2}^{k} b(j,k,p) < n^{-\frac{8}{7}}$.

Hence, when $n \to +\infty$, with probability more than $1 - n^{-\frac{8}{7}}$, $\ell_1' \leq \lceil \frac{1}{6}\log_s n\rceil$ and $L_{\ell_1} \geq (s-1)n^{\frac{1}{6}}/s \geq n^{\frac{1}{6}}/2$.

Stage 2: [27] proved that when $\widehat{L}_{i-1} \geq n^{\frac{1}{6}}/2$ and $\widehat{B}_{i-1} < n/s^4$,

$$\Pr\left[\widehat{L}_i \geq \left(1 - \frac{s+2}{s^4}\right) s\widehat{L}_{i-1} \mid \widehat{L}_{i-1}, \widehat{B}_{i-1}\right] > 1 - n^{-C}$$

for any fixed C and sufficiently large n. We then have that when $L_{i-1} \geq n^{\frac{1}{6}}/2$ and $B_{i-1} < n/s^4$,

$$\begin{aligned}&\Pr\left[L_i \geq \left(1 - \frac{s+2}{s^4}\right) sL_{i-1} \mid L_{i-1}, B_{i-1}\right]\\ >&\Pr\left[\widehat{L}_i \geq \left(1 - \frac{s+2}{s^4}\right) s\widehat{L}_{i-1} \mid \widehat{L}_{i-1} = L_{i-1}, \widehat{B}_{i-1} = B_{i-1}\right]\\ >&1 - n^{-C}\end{aligned}$$

for any fixed $C > 1$ and sufficiently large n. Thus, with probability $> \left(1 - n^{-C}\right)^{\ell_2'} > \left(1 - n^{-C}\right)^{n} > 1 - n^{1-C}$, all the levels constructed at Stage 2 have growth factor at least $s(1-(s+2)/s^4)$. With probability $> \left(1 - n^{-\frac{8}{7}}\right)\left(1 - n^{1-C}\right) \geq 1 - n^{-\frac{9}{8}}$,

$$\ell_2 < \log_{(1-(s+2)/s^4)s} n = \frac{1}{1 + \log_s(1-(s+2)/s^4)} \log_s n$$

and $B_{\ell_2} \geq n/s^4$ and for any $i \leq \ell_2$, $B_i > \left(\left(1-(s+2)/s^4\right)s\right)^i$.

Stage 3: [27] proved that for sufficiently large n,

$$\prod_{r=n/s^5}^{(1-2^{-s})n} \Pr\left[\widehat{B}(r) \geq r + C_s n\right] > 1 - n^{-C}$$

for a constant $C > 0$ and another constant C_s only depending on s. We then know

$$\prod_{r=n/s^5}^{(1-2^{-s})n} \Pr\left[B(r) \geq r + C_s n\right] > \prod_{r=n/s^5}^{(1-2^{-s})n} \Pr\left[\widehat{B}(r) \geq r + C_s n\right] > 1 - n^{-C}$$

for a constant $C > 0$ and another constant C_s only depending on s. This means that all the levels constructed at Stage 3 have at least $C_s n$ vertices with high probability. Formally, when $n \to +\infty$, with probability greater than $1 - n^{-C}$, $\ell'_3 < \frac{n}{C_s n} = \frac{1}{C_s}$.

So far, after Stage 3, there are only $n/2^s$ unoccupied vertices in the graph. If $s \geq \log_2 n - \log_2(C' \log_s n)$ for some constant $C' > 0$, we have $\frac{n}{2^s} \leq \frac{C' n \log_s n}{n} = O(\log_s n)$. That is, the number of unoccupied vertices is $O(\log_s n)$. No matter what will happen in Stage 4 to 6, in the worst case, the diameter of the graph will be at most $\ell'_1 + \ell'_2 + \ell'_3 + O(\log_s n) = O(\log_s n)$ and we are done.

However, if $s < \log_2 n - \log_2(C' \log_s n)$, we have to move on to the later stages.

Stage 4: We prove the boundary of the spanning ball starts shrinking in Stage 4.

$$\begin{aligned}
\Pr\left[L_i \leq \frac{1.5s}{2^s} L_{i-1} \mid L_{i-1}, B_{i-1}\right] &\geq \Pr\left[X\left(sL_{i-1}, \frac{n - B_{i-1}}{n-s}\right) \leq \frac{1.5s}{2^s} L_{i-1}\right] \\
&\geq \Pr\left[X\left(sL_{i-1}, \frac{n}{(n-s)2^s}\right) \leq \frac{1.5s}{2^s} L_{i-1}\right] \\
&= 1 - \Pr\left[X\left(sL_{i-1}, \frac{n}{(n-s)2^s}\right) > \frac{1.5s}{2^s} L_{i-1}\right] \\
&\geq 1 - sL_{i-1} \cdot \left(\frac{n-s}{n} \cdot 1.5\right)^{\left(\frac{n}{n-s} - 1.5\right)\left(\frac{sL_{i-1}}{2^{s+1}}\right) + \frac{3}{2}}
\end{aligned}$$

Because $s < \log_2 n - \log_2(C' \log_s n)$ and $L_{i-1} > (\log_2 n)^2$, it follows that when $n \to +\infty$, the above probability is at least $1 - n^{-C}$ for some constant $C > 1$. Formally, with probability at least $\left(1 - n^{-C}\right)^n > 1 - n^{1-C}$, all the levels constructed at Stage 4 have growth factor at most $\frac{1.5s}{2^s}$ and

$$\ell'_4 < \log_{2^s/(1.5s)} n = \frac{\log_2 s}{s - \log_2(1.5s)} \log_s n$$

Stage 5: We show the growth factor at this stage is at most $2/3$. Using the fact that $s \geq 2$, $L_{i-1} > 120 \log_2 n$ and $s < \log_2 n - \log_2(C' \log_s n)$, for sufficiently large n,

$$\begin{aligned}
\Pr\left[L_i \leq \frac{2}{3} L_{i-1} \mid L_{i-1}, B_{i-1}\right] &\geq 1 - \Pr\left[X\left(sL_{i-1}, \frac{n}{(n-s)2^s}\right) > \frac{2}{3} L_{i-1}\right] \\
&\geq 1 - sL_{i-1} \cdot \left(\frac{(n-s)2^{s+1}}{3ns}\right)^{\left(\frac{ns}{(n-s)2^{s+1}} - \frac{1}{3}\right) L_{i+1} + \frac{3}{2}} \\
&> 1 - sL_{i-1} \cdot 2^{1 - \frac{1}{30} L_{i-1}} \\
&> 1 - sL_{i-1} \cdot 2^{1 - 4 \log_2 n} \\
&> 1 - n^{-3}
\end{aligned}$$

This implies that all levels constructed at Stage 5 have growth at most 2/3 and $\ell_5' < \log_{\frac{3}{2}} (\log_2 n)^2 < (4 \log_2 s) \log_s \log_2 n$ with probability greater than $\left(1 - n^{-3}\right)^n > 1 - n^{-2}$.

Stage 6: We construct a logarithmic upper bound for the number of new vertices occupied at Stage 6. For some constant $C > 0$,

$$
\begin{aligned}
&\Pr\left[B_{\ell_6} - B_{\ell_5} > \frac{C \log_2 n}{s}\right] \\
\leq &\Pr\left[L_i > \frac{C \log_2 n}{s} \mid L_{i-1} = 120 \log_2 n + \frac{C \log_2 n}{s}, B_{i-1}\right] \\
\leq &\Pr\left[X\left(120 s \log_2 n + C \log_2 n, \frac{n}{(n-s)2^s}\right) > \frac{C \log_2 n}{s}\right] \\
\leq &(120s + C) \log_2 n \cdot \left(\frac{C(n-s)2^s}{ns(120s+C)}\right)^{\left((120s+C)2^{-s} - \frac{C}{s}\right) \log_2 \sqrt{n} + \frac{3}{2}} \\
\leq &(120s + C) \log_2 n \cdot 2^{\left(s + \log_2 \frac{C(n-s)}{ns(120s+C)}\right)\left((120s+C)2^{-s} - \frac{C}{s}\right) \log_2 \sqrt{n} + \frac{3}{2}\left(s + \log_2 \frac{C(n-s)}{ns(120s+C)}\right)}
\end{aligned}
$$

Simple algebra gives

$$
\begin{aligned}
&\left(s + \log_2 \frac{C(n-s)}{ns(120s+C)}\right)\left(\frac{120s+C}{2^s} - \frac{C}{s}\right) \\
= &-C + \frac{120s+C}{2^s} \log_2 \frac{C(n-s)}{ns(120s+C)} - \frac{C}{s} \log_2 \frac{C(n-s)}{ns(120s+C)} + \frac{120s^2 + Cs}{2^s}
\end{aligned}
$$

For any $s < \log_2 n - \log_2 (C' \log_s n)$, all the addends expect the first item approach zero as s increases. Therefore, there exists some constant C_0 such that when $n \to +\infty$, $\Pr\left[B_{\ell_6} - B_{\ell_5} > \frac{C_0 \log_2 n}{s}\right] < n^{-2}$. Formally, with probability greater than $1 - n^{-2}$,

$$
\ell_6' \leq B_{\ell_6} - B_{\ell_5} \leq \frac{C_0 \log_2 n}{s} = \frac{C_0 \log_2 s}{s} \log_s n
$$

Conclusion: With probability greater than $1 - n^{-\frac{10}{9}}$, the diameter of a RSG$^+(s)$ is at most $\sum_{i=1}^{6} \ell_i' = O(\log_s n)$. ■

With Theorem 3 and 4, we reach the fast convergence argument on the RSG$^+(s)$ model.

Appendix B Proof of Theorem 1 for random in-regular graphs and random s-in s-out graphs

The conclusions drawn on random walk on a random out-regular graph can be easily generalized to the in-regular cases. Let A be the adjacency matrix of graph G. Denote by $G^\top$ the *transpose* of G defined by adjacency matrix $A^\top$. The following facts are immediate observations from the definitions.

Fact 1 *For any $u, v \in V$, u and v are strongly connected in G if and only if they are strongly connected in $G^\top$.*

Fact 2 *Graph G is h-partite if and only if graph $G^\top$ is h-partite.*

Fact 3 *The distance from $u \in V$ to $v \in V$ in G is equal to the distance from v to u in $G^\top$.*

Fact 1 tells us $G^\top$ has exactly the same irreducible components as G and Fact 2 shows the equivalence of the aperiodicity of G and $G^\top$. Fact 3 leads to $Diam(G) = Diam(G^\top)$. Because a random in-regular graph can be created by transposing a corresponding random out-regular graph, we can conclude the following statements.

Corollary 1 *With probability $1 - o(1)$, a RMG$^-$(s) has a strongly connected component, denoted by $\tilde{G} = (\tilde{V}, \tilde{E})$, of size $\tilde{n}$ when $n \to +\infty$, and a) $\lim_{n\to+\infty} \frac{\tilde{n}}{n} = C$ for some constant $C > 0.7968$ when $s \geq 2$ or some $C > 0.999$ when $s \geq 7$; b) a random walk on $\tilde{G}$ is aperiodic.*

Corollary 2 *With probability $1-o(1)$, the diameter of a RMG$^-$(s) is $\Theta(\log_s n)$.*

Corollary 3 *With probability $1 - o(1)$, a RSG$^-$(s) has a strongly connected component, denoted by $\tilde{G} = (\tilde{V}, \tilde{E})$, of size $\tilde{n}$ when $n \to +\infty$, and a) $\lim_{n\to+\infty} \frac{\tilde{n}}{n} \geq C$ for some constant $C > 0.7968$ when $s \geq 2$ or some $C > 0.999$ when $s \geq 7$; b) a random walk on $\tilde{G}$ is aperiodic.*

Corollary 4 *With probability $1-o(1)$, the diameter of a RSG$^-$(s) is $\Theta(\log_s n)$.*

Note that in these cases the irreducible component is usually not closed. Hence, the fast convergence argument only holds when the walk is restricted to the unique irreducible component. According to Theorem 3, to bound the convergence rate we still need the maximum out-degree $s_0 = \arg\max_{u\in V} d_u$. To achieve fast convergence, we need a lower-bound assumption on the in-degree s.

Lemma 5 *Let $s_0 = \arg\max_{u\in V} d_u$ be the maximum out-degree of a RMG$^-$(s) with $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$ for some constant $C' \geq 1$. With probability $1-o(1)$, $s_0 = O(s^{C'+\varepsilon})$ for any constant $\varepsilon > 0$.*

Proof According to the properties of a RMG$^-(s)$, the probability of $s_0 > ns$ is 0 and $\Pr[s_0 = ns] \leq n \cdot n^{-ns}$ is exponentially small. For any $k < ns$,

$$\begin{aligned}
&\Pr[s_0 \geq k] \leq n \cdot \Pr[\text{a particular vertex has out-degree at least } k] \\
\leq & n \cdot \binom{ns}{k} \left(\frac{1}{n}\right)^k \\
\leq & \frac{\sqrt{2\pi ns}\left(\frac{ns}{e}\right)^{ns} e^{\frac{1}{12ns}}}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k e^{\frac{1}{12k+1}} \cdot \sqrt{2\pi(ns-k)}\left(\frac{ns-k}{e}\right)^{ns-k} e^{\frac{1}{12(ns-k)+1}}} \cdot n\left(\frac{1}{n}\right)^k \\
\leq & \sqrt{\frac{n^3 s}{2\pi k(ns-k)}} \cdot \frac{e^{\frac{1}{12ns}}(ns)^{ns}}{(nk)^k(ns-k)^{ns-k}} \\
\leq & \sqrt{\frac{1}{2\pi}} \cdot e^{\log n + ns\log(ns) - k\log k - (ns-k)\log(ns-k) - k\log n + \frac{1}{12ns}}
\end{aligned}$$

We only need to choose a k such that the exponent goes to $-\infty$ when $n \to +\infty$, which is equal to

$$\log n + k\left(1 - \frac{ns}{k}\right)\log\left(1 - \frac{k}{ns}\right) + k\log s - k\log k + \frac{1}{12ns}$$

Let $k = s^c$ where $c = C' + \varepsilon$. If $k \geq ns$ then $\Pr[s_0 \geq k]$ is exponentially small as discussed above. Otherwise we have $\left(1 - \frac{ns}{k}\right)\log\left(1 - \frac{k}{ns}\right) \leq 1$ in our case. Also notice that $\frac{1}{12ns} \leq 1$. The exponent is then upper bounded by $\log n + s^c - s^c(c-1)\log s + 1$. Letting $\log n \leq s^c(c - 1 - 0.5\varepsilon)\log s$ gives

$$s \geq \left[\frac{c\log n}{(c-1-0.5\varepsilon)W\left(\frac{c\log n}{c-1-0.5\varepsilon}\right)}\right]^{\frac{1}{c}} = o\left(\left[\frac{\log n}{\log\log n}\right]^{\frac{1}{C'}}\right)$$

where $W(x)$ is the Lambert W-function. ■

Combining Lemma 5 with Theorem 3, we reach the fast convergence of a random walk on a RMG$^-(s)$ with $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$.

The same convergence property holds on a RSG$^-(s)$.

Lemma 6 *Let* $s_0 = \arg\max_{u \in V} d_u$ *be the maximum out-degree of a RSG*$^-$*(s) with* $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$ *for some constant* $C' \geq 1$*. With probability* $1 - o(1)$, $s_0 = O(s^{C'+\varepsilon})$ *for any constant* $\varepsilon > 0$.

Proof From the definition of a RSG$^-(s)$, the probability of $s_0 \geq n$ is 0. If we have large $s = \Theta(n)$, then the argument automatically holds because $s_0 \leq n - 1 = O(s)$. Otherwise $s = o(n)$, $\Pr[s_0 = n-1] \leq n \cdot \left(\frac{s}{n-1}\right)^{n-1}$ is exponentially small. For any $k < n-1$, using the union bound,

$$\begin{aligned}
&\Pr[s_0 \geq k] \leq n \cdot \Pr[\text{a particular vertex has at least } k \text{ neighbors}]\\
\leq& n \cdot \binom{n-1}{k} \left[\frac{\binom{1}{1}\binom{n-2}{s-1}}{\binom{n-1}{s}} \right]^k\\
=& n \cdot \binom{n-1}{k} \left[\frac{s}{n-1} \right]^k\\
\leq& \frac{n \cdot \sqrt{2\pi(n-1)} \left(\frac{n-1}{e}\right)^{n-1} e^{\frac{1}{12(n-1)}}}{\sqrt{2\pi k} \left(\frac{k}{e}\right)^k e^{\frac{1}{12k+1}} \cdot \sqrt{2\pi(n-k-1)} \left(\frac{n-k-1}{e}\right)^{n-k-1} e^{\frac{1}{12(n-k-1)+1}}} \left(\frac{s}{n-1}\right)^k\\
\leq& \sqrt{\frac{n^2(n-1)}{2\pi k(n-k-1)}} \cdot \frac{e^{\frac{1}{12(n-1)}} (n-1)^{n-k-1} s^k}{k^k (n-k-1)^{n-k-1}}\\
\leq& \sqrt{\frac{1}{2\pi}} \cdot e^{\log n + \frac{1}{12(n-1)} + (n-k-1)\log(n-1) + k\log s - k\log k - (n-k-1)\log(n-k-1)}
\end{aligned}$$

Again we are supposed to set a proper value of k such that the exponent in the last expression goes to $-\infty$. The exponent can be reshaped as

$$\log n + \frac{1}{12(n-1)} + k\left(1 - \frac{n-1}{k}\right) \log\left(1 - \frac{k}{n-1}\right) + k\log s - k\log k$$

Because $\frac{1}{12(n-1)}$ and $\left(1 - \frac{n-1}{k}\right)\log\left(1 - \frac{k}{n-1}\right)$ are both at most 1 in our case, letting $c = C' + \varepsilon$ and $k = s^c$ gives us

$$\log n + 1 - s^c(c-1)\log s + s^c$$

For $s = \Omega\left(\left[\frac{\log n}{\log\log n}\right]^{1/C'}\right)$, the expression goes to $-\infty$ and completes the proof. ■

Thus we have proved the $\mathrm{RSG}^-(s)$ case in the main theorem.

The model of random s-in s-out graphs is a random graph model first introduced by Fenner and Frieze [12], which can be viewed as the sum of a random out-regular graph and a random in-regular graph, generated independently of each other. We provide a brief proof for $\mathrm{RMG}^{\pm}(s)$ by simply combining the previously proved arguments for $\mathrm{RMG}^+(s)$ and $\mathrm{RMG}^-(s)$. The same result for $\mathrm{RSG}^{\pm}(s)$ can be similarly achieved based on the arguments for $\mathrm{RSG}^+(s)$ and $\mathrm{RSG}^-(s)$.

The original paper by Fenner and Frieze [12] has already proved the strong connectivity of the random s-in s-out graphs for $s \geq 2$. As the entire graph is strongly connected, the connected component is surely closed and unique. As for aperiodicity, since $\tilde{V}$ is strongly connected, we only need to show one of the $v \in \tilde{V}$ is aperiodic. Without loss of generality, let $v \in \tilde{V}^+$ and then v is aperiodic in the RMG$^+(s)$ with high probability [1], which means that there exists a sufficiently large ℓ_0 such that for all $\ell \geq \ell_0$, there is a directed cycle of length ℓ over v. Because we only add edges onto the graph when generating the RMG$^-(s)$, the sum graph RMG$^\pm(s)$ still retains such cycles and v is aperiodic. The logarithmic diameter of RMG$^\pm(s)$ is due to

$$Diam(G_1 + G_2) \leq Diam(G_1) + Diam(G_2)$$

for any graphs G_1 and G_2.

Again, combining with Theorem 3 we reach the fast convergence property stated in the main theorem, and the same argument holds on a RSG$^\pm(s)$.

Appendix C Proof of Theorem 1 for random regular digraphs

In this section we continue showing positive results and study the random walks on RDG(s). Because the edges in this case are no longer chosen independently, the proof is done mainly by enumeration.

C.1 Irreducibility and aperiodicity

Previous works have shown the irreducibility [29].

Lemma 7 *With probability $1 - o(1)$, a RDG(s) is strongly connected when $s \geq 2$.*

Now we prove aperiodicity, starting with the asymptotic enumeration of regular digraphs.

Lemma 8 *Let $N(n, s)$ be the number of s-regular digraphs of size n.*

$$N(n,s) = \begin{cases} \frac{(ns)!}{(s!)^{2n}} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right] & \textit{if } 1 \leq s \leq \frac{n}{2} \\ N(n, n-s) & \textit{if } \frac{n}{2} < s < n \\ 1 & \textit{if } s = n \end{cases}$$

$N(n, s)$ is also the number of colored s-regular bipartite (undirected) graphs of size $2n$.

Proof We first show the bijection between regular digraphs of size n and colored regular bipartite (undirected) graphs of size $2n$. Let $G = (V, E)$ be a regular digraph of fixed degree s. We construct a regular bipartite graph $G' = (V', E')$ where $|V'| = 2|V|$ as following. Without loss of generality, denote $V = \{v_1, v_2, \ldots, v_n\}$ and $V' = \{v'_1, v'_2, \ldots, v'_{2n}\}$ with $\{v'_1, v'_2, \ldots, v'_n\}$ of one color and $\{v'_{n+1}, v'_{n+2}, \ldots, v'_{2n}\}$ of the other. Let $(v'_i, v'_{n+j}) \in E'$ if and only if $(v_i, v_j) \in E$. We can see such a regular bipartite graph G' is unique for each regular digraph G and vice versa. Note that this bijection is connectivity-preserving. To see this, consider that for a directed graph there are two cases of being disconnected. The first case is that there exist nonempty $V_1 \subset V$ and $V_2 \subset V$ with only edges going from V_1 to V_2 and no edge going back. This is impossible in a regular digraph because the in-degree of V_1 must be equal to its out-degree. The other case is no edge between V_1 and V_2, where the corresponding bipartite graph G' is also disconnected.

In order to prove the aperiodicity of a RDG(s), we first need to do enumeration for regular digraphs. It's easy to see another bijection: the one between regular digraphs and binary square matrices with equal line sums. Although little previous work has been done on the enumeration of regular digraphs, we are fortunate to have asymptotic results on the enumeration of binary square matrices with equal line sums. Let $N(n, s)$ be the number of regular digraphs with n vertices of fixed in-degree and out-degree equal to s, which is also the number of $n \times n$ binary matrices with equal line sums s and the number of regular bipartite graphs. McKay [19] proved that for $1 \leq s < \frac{1}{6}n$,

$$N(n,s) = \frac{(ns)!}{(s!)^{2n}} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right] \tag{1}$$

and Canfield and McKay [8] showed for $s \leq \frac{1}{2}n$ and $s = \Theta(n)$,

$$N(n,s) = \frac{\binom{n}{s}^{2n}}{\binom{n^2}{ns}} \left(1 - \frac{1}{n}\right)^{n-1} \exp\left(\frac{1}{2} + o(1)\right) \tag{2}$$

We are able to unify these two asymptotic results and show that the latter case also satisfies the former formula. For $s \leq \frac{1}{2}n$ and $s = \Theta(n)$,

$$\begin{aligned}
N(n,s) =& \frac{\binom{n}{s}^{2n}}{\binom{n^2}{ns}} \left(1 - \frac{1}{n}\right)^{n-1} \exp\left(\frac{1}{2} + o(1)\right) \\
=& \frac{\left(\frac{n!}{s!(n-s)!}\right)^{2n}}{\frac{(n^2)!}{(ns)!(n^2-ns)!}} \exp\left(o(1) - \frac{1}{2}\right) \\
=& \frac{(ns)!}{(s!)^{2n}} \cdot \frac{\left(\frac{n!}{(n-s)!}\right)^{2n}}{\frac{(n^2)!}{(n^2-ns)!}} \exp\left(O(1)\right) \\
=& \frac{(ns)!}{(s!)^{2n}} \cdot \frac{\left(\frac{\sqrt{2\pi n} \cdot n^n \cdot e^{n-s}}{e^n \cdot \sqrt{2\pi(n-s)} \cdot (n-s)^{n-s}}\right)^{2n}}{\frac{\sqrt{2\pi n^2} \cdot n^{2n^2} \cdot e^{n^2-ns}}{e^{n^2} \cdot \sqrt{2\pi(n^2-ns)} \cdot (n^2-ns)^{n^2-ns}}} \exp\left(O(1)\right) \\
=& \frac{(ns)!}{(s!)^{2n}} \cdot \frac{n^{2n^2} e^{ns} (n^2-ns)^{n^2-ns}}{e^{2ns} (n-s)^{2n(n-s)} n^{2n^2}} \cdot \left(\frac{n}{n-s}\right)^{n-\frac{1}{2}} \exp\left(O(1)\right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{(ns)!}{(s!)^{2n}} \cdot \frac{n^{n^2-ns}(n-s)^{n^2-ns}n^{n-\frac{1}{2}}}{e^{ns}(n-s)^{2n(n-s)}(n-s)^{n-\frac{1}{2}}} \exp\left(O(1)\right) \\
&= \frac{(ns)!}{(s!)^{2n}} \cdot \frac{n^{n^2-(s-1)n-\frac{1}{2}}}{e^{ns}(n-s)^{n^2-(s-1)n-\frac{1}{2}}} \exp\left(O(1)\right) \\
&= \frac{(ns)!}{(s!)^{2n}} \cdot \left(1-\frac{s}{n}\right)^{(s-1)n+\frac{1}{2}-n^2} \exp\left(O(1)-ns\right)
\end{aligned}
$$

Let $C = \frac{s}{n} \leq \frac{1}{2}$. Since $s = \Theta(n)$,

$$
\begin{aligned}
N(n,s) &= \frac{(ns)!}{(s!)^{2n}} \exp\left(\log(1-C) \cdot \left((s-1)n + \frac{1}{2} - n^2\right) + O(1) - ns\right) \\
&= \frac{(ns)!}{(s!)^{2n}} \exp\left((C-1)n^2 \log(1-C) - n\log(1-C) + O(1) - ns\right) \\
&= \frac{(ns)!}{(s!)^{2n}} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right]
\end{aligned}
$$

Using complement graphs, it is apparent that $N(n,s) = N(n, n-s)$ for $\frac{n}{2} \leq s < n$. When $s = n$, the only possible regular digraph in this case is the complete graph so $N(n,n) = 1$. Combining all the above cases completes the proof. ■

Lemma 9 *With probability* $1 - o(1)$*, a RDG(s) is aperiodic.*

Proof A regular digraph $G = (V, E)$ is h-partite if V can be partitioned into h subsets, $V_0, V_1, \ldots, V_{h-1}$, such that all edges from V_i go to $V_{(i+1) \mod h}$. Because the graph is regular, we must have $|V_0| = |V_1| = \ldots = |V_{h-1}| = \frac{n}{h}$ and $h \leq \frac{n}{s}$. Also we notice that the number of possible edge combinations from V_i going to $V_{(i+1) \mod h}$ is exactly the number of colored s-regular bipartite (undirected) graphs of size $\frac{n}{h}$, which is $N(\frac{n}{h}, s)$. Denote by p_h the probability of a RDG(s) being h-partite. The proof is done by showing p_h goes to 0 exponentially fast for all $2 \leq h \leq \frac{n}{s}$. The case where $s > \frac{n}{2}$ is trivial. Thus below we only consider $s \leq \frac{n}{2}$. We first prove the argument holds when $s = o(n)$. For $2 \leq h \leq \frac{n}{2s}$,

$$
p_h \leq \frac{1}{h} \cdot \binom{n}{\frac{n}{h}\ \frac{n}{h} \ldots \frac{n}{h}} \cdot \frac{\left[N\left(\frac{n}{h}, s\right)\right]^h}{N(n,s)}
$$

According to Lemma 8,

$$
\begin{aligned}
N(n,s) &= \frac{(ns)!}{(s!)^{2n}} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right] \\
&= \frac{\sqrt{2\pi ns}(ns)^{ns}e^{2ns}}{e^{ns}\left[\sqrt{2\pi s} \cdot s^s\right]^{2n}} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right] \\
&= \frac{\sqrt{ns}}{s^n(2\pi)^{n-\frac{1}{2}}} \cdot \left(\frac{en}{s}\right)^{ns} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
\left[N\left(\frac{n}{h},s\right)\right]^h =&\left[\frac{\sqrt{s\frac{n}{h}}}{s^{\frac{n}{h}}(2\pi)^{\frac{n}{h}-\frac{1}{2}}}\cdot\left(\frac{en}{hs}\right)^{s\frac{n}{h}}\exp\left[-\frac{(s-1)^2}{2}+O\left(\frac{s^3h}{n}\right)\right]\right]^h\\
=&\frac{\left(\frac{ns}{h}\right)^{\frac{1}{2}h}}{s^n(2\pi)^{n-\frac{h}{2}}}\cdot\left(\frac{en}{hs}\right)^{ns}\exp\left[-\frac{h(s-1)^2}{2}+O\left(\frac{s^3h^2}{n}\right)\right]
\end{aligned}
$$

Also,

$$
\begin{aligned}
\binom{n}{\frac{n}{h}\,\frac{n}{h}\cdots\frac{n}{h}} =&\frac{n!}{\left(\frac{n}{h}!\right)^h}\leq\frac{\sqrt{2\pi n}\cdot n^n e^{n+\frac{1}{12n}}}{e^n\left(\sqrt{2\pi\frac{n}{h}}\left(\frac{n}{h}\right)^{\frac{n}{h}}\right)^h}\\
=&\frac{\sqrt{2\pi n}\cdot h^n e^{\frac{1}{12n}}}{\left(\frac{2\pi n}{h}\right)^{\frac{1}{2}h}}=(2\pi n)^{\frac{1}{2}(1-h)}\cdot h^{n+\frac{1}{2}h}\cdot e^{\frac{1}{12n}}
\end{aligned}
$$

We then have

$$
\begin{aligned}
p_h\leq&\frac{h^{n+\frac{1}{2}h-1}}{(2\pi n)^{\frac{1}{2}(h-1)}}\cdot\frac{\left(\frac{ns}{h}\right)^{\frac{1}{2}h}s^n(2\pi)^{n-\frac{1}{2}}}{s^n(2\pi)^{n-\frac{1}{2}h}\sqrt{ns}}\cdot\left(\frac{1}{h}\right)^{ns}\\
&\cdot\exp\left[\frac{(s-1)^2}{2}-\frac{h(s-1)^2}{2}+O\left(\frac{s^3h^2}{n}\right)\right]\\
=&\frac{h^{n+\frac{1}{2}h-1}}{(2\pi n)^{\frac{1}{2}(h-1)}}\cdot\frac{(2\pi ns)^{\frac{1}{2}(h-1)}}{h^{\frac{1}{2}h}}\cdot\left(\frac{1}{h}\right)^{ns}\\
&\cdot\exp\left[-\frac{(h-1)(s-1)^2}{2}+O\left(\frac{s^3h^2}{n}\right)\right]\\
=&\frac{s^{\frac{1}{2}(h-1)}}{h^{ns-n+1}}\exp\left[-\frac{(h-1)(s-1)^2}{2}+O\left(\frac{s^3h^2}{n}\right)\right]\\
=&\exp\left\{-\frac{1}{2}(h-1)[(s-1)^2-\log s]-(ns-n+1)\log h+O\left(\frac{s^3h^2}{n}\right)\right\}
\end{aligned}
$$

For $s\geq 2$, we have $(s-1)^2-\log s>0$. When $h=O(1)$, $O\left(\frac{s^3h^2}{n}\right)=O\left(\frac{s^3}{n}\right)=o(ns)$ since $s=o(n)$. When $h=\omega(1)$, as $h<\frac{n}{s}$, $O\left(\frac{s^3h^2}{n}\right)=O(ns)=o(ns\log h)$. Hence, p_h goes to 0 exponentially fast.

For $\frac{n}{2s} < h < \frac{n}{s}$, surely $h = \omega(1)$ as $s = o(n)$. Also, $\left(\frac{n}{h} - s\right)^3 < s^3$ since $\frac{n}{2h} < s < \frac{n}{h}$. Then we have

$$\begin{aligned}
\left[N\left(\frac{n}{h}, \frac{n}{h} - s\right)\right]^h =& \Bigg[\frac{\sqrt{\frac{n}{h}\left(\frac{n}{h} - s\right)}}{\left(\frac{n}{h} - s\right)^{\frac{n}{h}} (2\pi)^{\frac{n}{h} - \frac{1}{2}}} \cdot \left(\frac{en}{h\left(\frac{n}{h} - s\right)}\right)^{\frac{n}{h}\left(\frac{n}{h} - s\right)} \\
& \cdot \exp\left[-\frac{\left(\frac{n}{h} - s - 1\right)^2}{2} + O\left(\frac{\left(\frac{n}{h} - s\right)^3 h}{n}\right)\right]\Bigg]^h \\
=& \frac{\left(\frac{n}{h}\left(\frac{n}{h} - s\right)\right)^{\frac{1}{2}h}}{\left(\frac{n}{h} - s\right)^n (2\pi)^{n - \frac{h}{2}}} \cdot \left(\frac{en}{h\left(\frac{n}{h} - s\right)}\right)^{n\left(\frac{n}{h} - s\right)} \\
& \cdot \exp\left[-\frac{h\left(\frac{n}{h} - s - 1\right)^2}{2} + O\left(\frac{s^3 h^2}{n}\right)\right]
\end{aligned}$$

so that

$$\begin{aligned}
p_h \le& \frac{1}{h} \cdot \binom{n}{\frac{n}{h}\,\frac{n}{h}\,\ldots\,\frac{n}{h}} \cdot \frac{\left[N(\frac{n}{h}, \frac{n}{h} - s)\right]^h}{N(n, s)} \\
=& \frac{h^{n + \frac{1}{2}h - 1}}{(2\pi n)^{\frac{1}{2}(h-1)}} \cdot \frac{\left(\frac{n}{h}\left(\frac{n}{h} - s\right)\right)^{\frac{1}{2}h} s^n (2\pi)^{n - \frac{1}{2}}}{\left(\frac{n}{h} - s\right)^n (2\pi)^{n - \frac{1}{2}h} \sqrt{ns}} \cdot \left(\frac{en}{h\left(\frac{n}{h} - s\right)}\right)^{n\left(\frac{n}{h} - s\right)} \\
& \cdot \left(\frac{en}{s}\right)^{-ns} \cdot \exp\left[\frac{(s-1)^2}{2} - \frac{h\left(\frac{n}{h} - s - 1\right)^2}{2} + O\left(\frac{s^3 h^2}{n}\right)\right]
\end{aligned}$$

Notice that function $\left(\frac{ey}{x}\right)^x$ with constraint $0 < x \le \frac{1}{2}y$ reaches its maximum at $x = \frac{1}{2}y$, which implies

$$\left(\frac{en}{h\left(\frac{n}{h} - s\right)}\right)^{\frac{n}{h} - s} \le (2e)^{\frac{n}{2h}}$$

and

$$\begin{aligned}
p_h \leq & h^{n+\frac{1}{2}h-1} \cdot \frac{\sqrt{2\pi n}}{\sqrt{ns}} \left(\frac{n}{2\pi h n}\right)^{\frac{1}{2}h} \cdot \frac{(2\pi)^{\frac{1}{2}(h-1)}}{\left(\frac{n}{h}-s\right)^{n-\frac{1}{2}h}} \cdot s^n \cdot \left(\frac{2s}{n}\right)^{ns} \\
& \cdot (2e)^{\frac{n^2}{2h}-ns} \cdot \exp\left[\frac{(s-1)^2}{2} - \frac{h\left(\frac{n}{h}-s-1\right)^2}{2} + O\left(\frac{s^3h^2}{n}\right)\right] \\
= & s^{n-\frac{1}{2}} \left(\frac{n}{h}-s\right)^{\frac{1}{2}h-n} \cdot h^{n-1}(2e)^{\frac{n^2}{2h}-ns} \left(\frac{2s}{n}\right)^{ns} \\
& \cdot \exp\left[\frac{(s-1)^2}{2} - \frac{h\left(\frac{n}{h}-s-1\right)^2}{2} + O\left(\frac{s^3h^2}{n}\right)\right] \\
= & s^{-\frac{1}{2}} h^{-1} \left(\frac{n}{h}-s\right)^{\frac{1}{2}h-n} \cdot (2e)^{\frac{n^2}{2h}-ns} \cdot \exp\Bigg[n\log s + n\log h + ns\log 2 + ns\log s \\
& - ns\log n - \frac{n^2}{2h} - \frac{hs^2}{2} - hs - \frac{h}{2} + ns + n + \frac{1}{2}s^2 - s + O\left(\frac{s^3h^2}{n}\right)\Bigg] \\
= & s^{-\frac{1}{2}} h^{-1} \left(\frac{n}{h}-s\right)^{\frac{1}{2}h-n} \cdot (2e)^{\frac{n^2}{2h}-ns} \cdot \exp\Bigg[n\log s + ns\log 2 + ns\log s \\
& - \left(1 - \frac{\log h}{s\log n}\right) ns\log n - \frac{n^2}{2h} - \frac{hs^2}{2} - hs - \frac{h}{2} + ns + n + \frac{1}{2}s^2 - s + O\left(\frac{s^3h^2}{n}\right)\Bigg]
\end{aligned}$$

Notice that $O\left(\frac{s^3h^2}{n}\right) = O(ns)$ for $h < \frac{n}{s}$ and $\frac{n^2}{2h} - ns < 0$ for $s > \frac{n}{2h}$. Also, $1 - \frac{\log h}{s\log n} > 0$ for any $s \geq 2$, we have p_h going to 0 exponentially fast.

The case where $h = \frac{n}{s}$ is deferred to the end of this proof.

Now we study the case when $s = \Theta(n) < \frac{1}{2}n$ and $\frac{n}{h} - s = \Theta(n)$. In this case we have $\epsilon n \leq s \leq \left(\frac{1}{h} - \epsilon\right) n$ for some positive constant $\epsilon > 0$ and $2 \leq h < \frac{n}{s}$ is surely $O(1)$. Let $C = \frac{s}{n} < \frac{1}{2}$ so $0 < \epsilon \leq \lim_{n\to+\infty} C \leq \frac{1}{h} - \epsilon < \frac{1}{2}$. When $2 \leq h \leq \frac{n}{2s}$, we have $s = \Theta(n) = \Theta(\frac{n}{h})$.

$$p_h \leq \frac{1}{h} \cdot \binom{n}{\frac{n}{h}\ \frac{n}{h} \ldots \frac{n}{h}} \cdot \frac{\left[N\left(\frac{n}{h}, s\right)\right]^h}{N(n,s)}$$

According to Lemma 8,

$$\begin{aligned}
N(n,s) = & \frac{(ns)!}{(s!)^{2n}} \exp\big((C-1)n^2\log(1-C) - n\log(1-C) + O(1) - ns\big) \\
= & \frac{\sqrt{2\pi ns}(ns)^{ns}e^{2ns}}{e^{ns}\left[\sqrt{2\pi s}\cdot s^s\right]^{2n}} \exp\big((C-1)n^2\log(1-C) - n\log(1-C) + O(1) - ns\big) \\
= & \frac{\sqrt{ns}}{s^n(2\pi)^{n-\frac{1}{2}}} \cdot \left(\frac{n}{s}\right)^{ns} \exp\big((C-1)n^2\log(1-C) - n\log(1-C) + O(1)\big)
\end{aligned}$$

and

$$\begin{aligned}\left[N\left(\frac{n}{h},s\right)\right]^h =&\left[\frac{\sqrt{s\frac{n}{h}}}{s^{\frac{n}{h}}(2\pi)^{\frac{n}{h}-\frac{1}{2}}}\cdot\left(\frac{n}{hs}\right)^{s\frac{n}{h}}\exp\left((hC-1)\left(\frac{n}{h}\right)^2\log(1-hC)-\log(1-hC)\cdot\frac{n}{h}+O(1)\right)\right]^h\\ =&\frac{\left(\frac{ns}{h}\right)^{\frac{1}{2}h}}{s^n(2\pi)^{n-\frac{h}{2}}}\cdot\left(\frac{n}{hs}\right)^{ns}\exp\left((hC-1)\frac{n^2}{h}\log(1-hC)-n\log(1-hC)+O(h)\right)\end{aligned}$$

so that

$$\begin{aligned}p_h \le&\frac{h^{n+\frac{1}{2}h-1}}{(2\pi n)^{\frac{1}{2}(h-1)}}\cdot\frac{\left(\frac{ns}{h}\right)^{\frac{1}{2}h}s^n(2\pi)^{n-\frac{1}{2}}}{s^n(2\pi)^{n-\frac{1}{2}h}\sqrt{ns}}\cdot\left(\frac{1}{h}\right)^{ns}\cdot\exp\Bigg[(hC-1)\frac{n^2}{h}\log(1-hC)\\ &-n\log(1-hC)+O(h)-(C-1)n^2\log(1-C)+n\log(1-C)-O(1)\Bigg]\\ =&\frac{h^{n+\frac{1}{2}h-1}}{(2\pi n)^{\frac{1}{2}(h-1)}}\cdot\frac{(2\pi ns)^{\frac{1}{2}(h-1)}}{h^{\frac{1}{2}h}}\cdot\left(\frac{1}{h}\right)^{ns}\cdot\exp\Bigg[\left(\left(C-\frac{1}{h}\right)\log(1-hC)-(C-1)\log(1-C)\right)n^2\\ &+(\log(1-C)-\log(1-hC))n+O(h)\Bigg]\\ =&\frac{s^{\frac{1}{2}(h-1)}}{h^{ns-n+1}}\cdot\exp\Bigg[\left(\left(C-\frac{1}{h}\right)\log(1-hC)-(C-1)\log(1-C)\right)n^2\\ &+(\log(1-C)-\log(1-hC))n+O(h)\Bigg]\end{aligned}$$

Notice that function $\left(y-\frac{1}{x}\right)\log(1-xy)$ with constraints $xy\le\frac{1}{2}$ and $x\ge 2$ reaches its maximum at $x=2$. Thus

$$\left(C-\frac{1}{h}\right)\log(1-hC)\le\left(C-\frac{1}{2}\right)\log(1-2C)$$

Also note that function $f(x)=(1-x)\log(1-x)+\left(x-\frac{1}{2}\right)\log(1-2x)<0$ for any $0<x<\frac{1}{2}$. Therefore, p_h goes to 0 exponentially fast.

When $\frac{n}{2s} < h < \frac{n}{s}$, as $\frac{n}{h} - s = \Theta(n) = \Theta(\frac{n}{h})$,

$$\begin{aligned}\left[N\left(\frac{n}{h}, \frac{n}{h} - s\right)\right]^h =& \Bigg[\frac{\sqrt{\frac{n}{h}\left(\frac{n}{h} - s\right)}}{\left(\frac{n}{h} - s\right)^{\frac{n}{h}}(2\pi)^{\frac{n}{h} - \frac{1}{2}}} \cdot \left(\frac{n}{h\left(\frac{n}{h} - s\right)}\right)^{\frac{n}{h}\left(\frac{n}{h} - s\right)} \cdot \exp\Bigg(O(1) \\ &+ (1 - hC - 1)\left(\frac{n}{h}\right)^2 \log(1 - (1 - hC)) - \frac{n}{h}\log(1 - (1 - hC))\Bigg)\Bigg]^h \\ =& \frac{\left(\frac{n}{h}\left(\frac{n}{h} - s\right)\right)^{\frac{1}{2}h}}{\left(\frac{n}{h} - s\right)^n (2\pi)^{n - \frac{h}{2}}} \cdot \left(\frac{n}{h\left(\frac{n}{h} - s\right)}\right)^{n\left(\frac{n}{h} - s\right)} \\ &\cdot \exp\Bigg(-Cn^2\log(hC) - n\log(hC) + O(h)\Bigg)\end{aligned}$$

and

$$\begin{aligned}p_h \le& \frac{1}{h} \cdot \binom{n}{\frac{n}{h}\,\frac{n}{h}\ldots\frac{n}{h}} \cdot \frac{\left[N(\frac{n}{h}, \frac{n}{h} - s)\right]^h}{N(n,s)} \\ =& \frac{h^{n + \frac{1}{2}h - 1}}{(2\pi n)^{\frac{1}{2}(h-1)}} \cdot \frac{\left(\frac{n}{h}\left(\frac{n}{h} - s\right)\right)^{\frac{1}{2}h} s^n (2\pi)^{n - \frac{1}{2}}}{\left(\frac{n}{h} - s\right)^n (2\pi)^{n - \frac{1}{2}h}\sqrt{ns}} \cdot \left(\frac{n}{h\left(\frac{n}{h} - s\right)}\right)^{n\left(\frac{n}{h} - s\right)} \\ &\cdot \left(\frac{n}{s}\right)^{-ns} \cdot \exp\Bigg(-Cn^2\log(hC) - n\log(hC) + O(h) \\ &- (C-1)n^2\log(1 - C) + n\log(1 - C) - O(1)\Bigg) \\ =& h^{n-1}s^{n - \frac{1}{2}} \cdot \left(\frac{n}{h} - s\right)^{\frac{1}{2}h - n} \cdot (1 - hC)^{\left(C - \frac{1}{h}\right)n^2} \cdot C^{Cn^2} \\ &\exp\Bigg((-C\log(hC) - (C-1)\log(1 - C))n^2 + (\log(1 - C) - \log(hC))n + O(h)\Bigg) \\ =& \frac{h^{n-1}}{\sqrt{s}} \cdot \left(\frac{n}{h} - s\right)^{\frac{1}{2}h - n} \cdot \exp\Bigg(\Bigg(\left(C - \frac{1}{h}\right)\log(1 - hC) + C\log C \\ &- C\log(hC) - (C-1)\log(1 - C)\Bigg)n^2 + n\log n + (\log(1 - C) - \log h)n + O(h)\Bigg)\end{aligned}$$

where

$$\begin{aligned}&\left(C - \frac{1}{h}\right)\log(1 - hC) + C\log C - C\log(hC) - (C-1)\log(1 - C) \\ =& \left(C - \frac{1}{h}\right)\log(1 - hC) - C\log h - (C-1)\log(1 - C)\end{aligned}$$

Notice that

$$\frac{\partial}{\partial h}\left[\left(C-\frac{1}{h}\right)\log\left(1-hC\right)-C\log h\right]=\frac{\log(1-Ch)}{h^2}<0$$

as $0<Ch=\frac{sh}{n}<1$. Due to $\frac{1}{2C}=\frac{n}{2s}<h<\frac{n}{s}=\frac{1}{C}$,

$$\begin{aligned}
&\left(C-\frac{1}{h}\right)\log\left(1-hC\right)+C\log C-C\log(hC)-(C-1)\log(1-C)\\
\leq&(C-2C)\log\left(1-C\frac{1}{2C}\right)-C\log\frac{1}{2C}-(C-1)\log(1-C)\\
=&C\log 2+C\log(2C)-(C-1)\log(1-C)\\
=&C\log(4C)-(C-1)\log(1-C)<0
\end{aligned}$$

for any $0<\epsilon\leq C\leq\frac{1}{2}-\epsilon<\frac{1}{2}$. Therefore, we again have p_h going to 0 exponentially fast.

When $s=\Theta(n)<\frac{n}{2}$ and $\frac{n}{h}-s=o(n)$, let $C=\frac{s}{n}<\frac{1}{h}$ but $\lim_{n\to+\infty}C=\frac{1}{h}$. In this case $2\leq h<\frac{n}{s}$ is $O(1)$ and surely $\frac{n}{2s}<h<\frac{n}{s}$. Otherwise, $h\leq\frac{n}{2s}$ implies $\frac{n}{h}-s\geq\frac{n}{2h}=\Theta(n)$. Also, due to $\frac{n}{h}-s=o(n)$ and $h=O(1)$, $O\left(\frac{\left(\frac{n}{h}-s\right)^3h^2}{n}\right)=o(n^2)$.

$$\begin{aligned}
p_h\leq&\frac{1}{h}\cdot\binom{n}{\frac{n}{h}\,\frac{n}{h}\,\ldots\,\frac{n}{h}}\cdot\frac{\left[N(\frac{n}{h},\frac{n}{h}-s)\right]^h}{N(n,s)}\\
=&\frac{h^{n+\frac{1}{2}h-1}}{(2\pi n)^{\frac{1}{2}(h-1)}}\cdot\frac{\left(\frac{n}{h}\left(\frac{n}{h}-s\right)\right)^{\frac{1}{2}h}s^n(2\pi)^{n-\frac{1}{2}}}{\left(\frac{n}{h}-s\right)^n(2\pi)^{n-\frac{1}{2}h}\sqrt{ns}}\cdot\left(\frac{en}{h\left(\frac{n}{h}-s\right)}\right)^{n\left(\frac{n}{h}-s\right)}\\
&\cdot\left(\frac{n}{s}\right)^{-ns}\cdot\exp\Bigg(-\frac{h\left(\frac{n}{h}-s-1\right)^2}{2}+O\left(\frac{\left(\frac{n}{h}-s\right)^3h^2}{n}\right)\\
&-(C-1)n^2\log(1-C)+n\log(1-C)-O(1)\Bigg)\\
=&h^{n-1}s^{n-\frac{1}{2}}\cdot\left(\frac{n}{h}-s\right)^{\frac{1}{2}h-n}\cdot\left(\frac{1-hC}{e}\right)^{\left(C-\frac{1}{h}\right)n^2}\cdot C^{Cn^2}\exp\Bigg(-\frac{n^2}{2h}-\frac{hs^2}{2}\\
&-hs-\frac{h}{2}+ns+n-(C-1)n^2\log(1-C)+n\log(1-C)+o(n^2)\Bigg)\\
=&\frac{h^{n-1}}{\sqrt{s}}\cdot\left(\frac{n}{h}-s\right)^{\frac{1}{2}h-n}\cdot\exp\Bigg(\Bigg(\left(\frac{1}{h}-C\right)-\left(\frac{1}{h}-C\right)\log\left(1-hC\right)\\
&+C\log C-(C-1)\log(1-C)-\frac{1}{2h}-\frac{hC^2}{2}+C\Bigg)n^2+o(n^2)\Bigg)
\end{aligned}$$

where $\lim_{n\to+\infty} C = \frac{1}{h}$ and

$$\begin{aligned}
&\left(\frac{1}{h}-C\right)-\left(\frac{1}{h}-C\right)\log\left(1-hC\right)+C\log C-(C-1)\log(1-C)-\frac{1}{2h}-\frac{hC^2}{2}+C\\
=&0+0-\frac{1}{h}\log h-\left(1-\frac{1}{h}\right)\log\left(\frac{h}{h-1}\right)-\frac{1}{2h}-\frac{1}{2h}+\frac{1}{h}\\
=&-\frac{1}{h}\log h-\left(1-\frac{1}{h}\right)\log\left(\frac{h}{h-1}\right)<0
\end{aligned}$$

for which p_h goes to 0 exponentially fast.

There are still two cases to be addressed, where $h=\frac{n}{s}$ or $s=\frac{1}{2}n$. Notice that when $s=\frac{1}{2}n$, the only possible h is $2=\frac{n}{s}$ so we can handle both by handling the former. In this case, we have

$$p_{\frac{n}{s}} \leq \frac{s}{n}\binom{n}{s\,s\ldots s}\cdot\frac{1}{N(n,s)}$$

where

$$\binom{n}{s\,s\ldots s}=\frac{n!}{(s!)^{\frac{n}{s}}}\leq\frac{\sqrt{2\pi n}\cdot n^n e^{n+\frac{1}{12n}}}{e^n\left(\sqrt{2\pi s}\cdot s^s\right)^{\frac{n}{s}}}=\frac{\sqrt{2\pi n}}{(2\pi s)^{\frac{n}{2s}}}\left(\frac{n}{s}\right)^n e^{\frac{1}{12n}}$$

For $s=o(n)$,

$$\begin{aligned}
p_{\frac{n}{s}} \leq&\frac{\sqrt{2\pi n}}{(2\pi s)^{\frac{n}{2s}}}\left(\frac{n}{s}\right)^{n-1}e^{\frac{1}{12n}}\cdot\frac{s^n(2\pi)^{n-\frac{1}{2}}}{\sqrt{ns}}\left(\frac{en}{s}\right)^{-ns}\exp\left(\frac{(s-1)^2}{2}\right)\\
=&\frac{(2\pi)^n}{\sqrt{s}(2\pi s)^{\frac{n}{2s}}}\left(\frac{s}{n}\right)^{ns-n+1}\exp\left(n\log s-ns+\frac{(s-1)^2}{2}+\frac{1}{12n}\right)\\
\leq&\frac{1}{\sqrt{s}(2\pi s)^{\frac{n}{2s}}}\left(\frac{2\pi s}{n}\right)^{n+1}\exp\left(n\log s-ns+\frac{(s-1)^2}{2}+\frac{1}{12n}\right)
\end{aligned}$$

which goes to 0 exponentially fast.

For $s=\Theta(n)$ and $s\leq\frac{1}{2}n$,

$$\begin{aligned}
p_{\frac{n}{s}} \leq&\frac{\sqrt{2\pi n}}{(2\pi s)^{\frac{n}{2s}}}\left(\frac{n}{s}\right)^{n-1}e^{\frac{1}{12n}}\cdot\frac{s^n(2\pi)^{n-\frac{1}{2}}}{\sqrt{ns}}\cdot\left(\frac{n}{s}\right)^{-ns}\\
&\exp(-(C-1)n^2\log(1-C)+n\log(1-C)-O(1))\\
=&\frac{(2\pi)^n}{\sqrt{s}(2\pi s)^{\frac{n}{2s}}}\left(\frac{s}{n}\right)^{ns-n+1}\exp\Big(-(C-1)n^2\log(1-C)\\
&+n\log n+(\log(1-C)+\log C)n+o(1)\Big)
\end{aligned}$$

which goes to 0 exponentially fast as well. Combining all the above cases completes the proof. ■

C.2 Fast convergence

According to Theorem 3, the proof of fast convergence can be done by either bounding the diameter of the graph or directly bounding the first non-zero eigenvalue of the Laplacian matrix. In this section we present the fast convergence of random walks on RSG$^+(s)$ via the spectral method.

First note that the walk matrix P of a random walk on a RDG(s) is doubly stochastic matrix, so is $\frac{1}{2}(P + P^\top)$. Fiedler [13] proved a very useful theorem:

Theorem 5 *Let Q be a doubly stochastic $n \times n$ matrix ($n \geq 2$) and $\lambda \neq 1$ be any non-stochastic eigenvalue of Q.*

$$|1 - \lambda| \geq \varphi_n[\mu(Q)]$$

where

$$\mu(Q) = \min_{\emptyset \neq M \subset [n]} \sum_{i \in M, j \notin M} Q_{ij}$$

and

$$\varphi_n(x) = \begin{cases} 2\left(1 - \cos\frac{\pi}{n}\right)x & \textit{if } 0 \leq x \leq \frac{1}{2} \\ 1 - 2(1-x)\cos\frac{\pi}{n} - (2x-1)\cos\frac{2\pi}{n} & \textit{if } \frac{1}{2} < x \leq 1 \end{cases}$$

The same paper also presented the following lemma.

Lemma 10 *For any doubly stochastic matrix Q, $0 \leq \mu(Q) \leq 1$. Q is reducible if and only if $\mu(Q) = 0$.*

Now we show the fast convergence of random walks on RDG(s).

Theorem 6 *With probability $1 - o(1)$, a random walk on a RDG(s) has $\Delta_{\chi^2}(t) \leq e^{-k}$ after at most $t \geq 2s(n-1)(\log n + 2k)$ steps.*

Proof As P has been shown irreducible with probability $1 - o(1)$, so is $\frac{1}{2}(P + P^\top)$. Then for Lemma 10 $0 < \mu(\frac{1}{2}(P + P^\top)) \leq 1$. The fact that any non-zero entry in P is at least $\frac{1}{s}$ gives $\mu(\frac{1}{2}(P + P^\top)) \geq \frac{1}{2s}$. For Theorem 5,

$$\left|1 - \lambda_{\frac{1}{2}(P+P^\top)}\right| \geq 2\left(1 - \cos\frac{\pi}{n}\right)\frac{1}{2s} > \frac{1}{s(n-1)}$$

for all non-stochastic eigenvalues $\lambda_{\frac{1}{2}(P+P^\top)} \neq 1$ of matrix $\frac{1}{2}(P + P^\top)$, due to the fact $\cos x < 1 - \frac{x}{\pi - x}$ for all $x \in \left(0, \frac{\pi}{2}\right)$. Also observing that the stationary distribution on a RDG(s) is always the uniform distribution, we have the Laplacian matrix

$$\mathcal{L} = I - \frac{\Phi^{\frac{1}{2}} P \Phi^{-\frac{1}{2}} + \Phi^{-\frac{1}{2}} P^\top \Phi^{\frac{1}{2}}}{2} = I - \frac{1}{2}(P + P^\top)$$

and $|\lambda_1(\mathcal{L})| \geq \left|1 - \lambda_{\frac{1}{2}(P+P^\top)}\right| > \frac{1}{s(n-1)}$ where $\lambda_1(\mathcal{L})$ is the smallest nonzero eigenvalue of $\mathcal{L}$. Combining with $\phi(u) = \frac{1}{n}$ for any $u \in V$ we complete the proof. ■

Appendix D Proof of Theorem 1 for random regular undirected graphs

Random regular undirected graphs are much more widely studied than directed ones, mainly because of the symmetry of undirected graphs. However, the study of the convergence of random walks on RG(s) is still very limited. Hildebrand [16] proved fast convergence with constraint $s = \lfloor \log^C n \rfloor$ for some constant $C \geq 2$. Cooper and Frieze [10] studied the cover time of RG(s) with fixed constant $s = O(1)$ but no convergence result was provided. In this section we present a more general result with constraint $3 \leq s = o(\sqrt{n})$ or $s > \frac{1}{2}n$. This constraint comes from the enumeration of RG(s) and the proof could be generalized if we have better results on the enumeration problem in the future.

Cooper et al. [11] and Krivelevich et al. [18] together proved the connectivity of RG(s) for $s \geq 3$.

Lemma 11 *With probability* $1 - o(1)$*, a RG(s) is connected when* $s \geq 3$*.*

Now we prove the aperiodicity as below.

Lemma 12 *With probability* $1 - o(1)$*, a RG(s) is aperiodic when* $s \geq 3$ *for odd* n*;* $3 \leq s = o(\sqrt{n})$ *or* $s > \frac{1}{2}n$ *for even* n*.*

Proof When n is odd, the graph is surely aperiodic because for undirected graphs the only periodic case is being bipartite and for regular undirected graphs the only bipartite partition is an even partition. Also, the aperiodicity is trivial when $s > \frac{1}{2}n$. Below we will prove the nontrivial case where n is even and $3 \leq s \leq \frac{1}{2}n$. Denote by $N'(n, s)$ the number of s-regular undirected graphs of size n. McKay and Wormald [20] proved an enumeration result for $s = o(\sqrt{n})$ that

$$N'(n, s) = \frac{(sn)!}{\left(\frac{1}{2}sn\right)! \cdot 2^{\frac{1}{2}ns}(s!)^n} \exp\left[\frac{1-s^2}{4} - \frac{s^3}{12n} + O\left(\frac{s^2}{n}\right)\right]$$

Since $s = o(\sqrt{n}) < \frac{1}{4}n$, the probability of a RG(s) being periodic p_2 is bounded by

$$\begin{aligned}
p_2 \leq & \frac{1}{2}\binom{n}{\frac{n}{2}} \cdot \frac{N(\frac{n}{2}, s)}{N'(n, s)} \\
= & \frac{1}{2}\frac{n!}{\left(\frac{n}{2}!\right)^2} \cdot \frac{\left(\frac{ns}{2}\right)!\left(\frac{ns}{2}\right)! \cdot 2^{\frac{ns}{2}}(s!)^n}{(s!)^n \cdot (ns)!} \exp\left[-\frac{(s-1)^2}{2} + O\left(\frac{s^3}{n}\right) + \frac{s^2}{4} - \frac{1}{4}\right] \\
= & \frac{\sqrt{2\pi n} \cdot n^n e^n}{2 \cdot e^n \pi n \left(\frac{n}{2}\right)^n} \cdot \frac{\left[\left(\frac{ns}{2}\right)!\right]^2 \cdot 2^{\frac{ns}{2}}}{(ns)!} \exp\left[-\frac{1}{4}s^2 + s + O\left(\frac{s^3}{n}\right) - \frac{3}{4}\right] \\
= & \frac{2^{n+\frac{1}{2}ns}}{\sqrt{2\pi n}} \cdot \frac{\pi ns \cdot \left(\frac{1}{2}ns\right)^{ns} e^{ns}}{e^{ns}(ns)^{ns}\sqrt{2\pi ns}} \exp\left[-\frac{1}{4}s^2 + s + O\left(\frac{s^3}{n}\right) - \frac{3}{4}\right] \\
= & 2^{-\frac{1}{2}ns+n-1} \cdot \sqrt{s} \cdot \exp\left[-\frac{1}{4}s^2 + s + O\left(\frac{s^3}{n}\right) - \frac{3}{4}\right]
\end{aligned}$$

When $s = \omega(1)$ and $s = o(\sqrt{n})$, p_2 goes to 0 exponentially fast because $O\left(\frac{s^3}{n}\right) = o(s)$. When $s = O(1)$ and $s \geq 3$, $-\frac{1}{2}s + 1 < 0$ and p_2 goes to 0 exponentially fast as well, which completes the proof. ■

The fast convergence argument for RG(s) can be proved using the same proof for RDG(s). The only difference is P being symmetric and $\frac{1}{2}(P + P^\top) = P$ so $|\lambda_1(\mathcal{L})| \geq |1 - \lambda_P| > \frac{2}{s(n-1)}$.

Appendix E Proof of Theorem 2 for RSG$^+(s)$

Lemma 13 *With probability* $1 - o(1)$*, a RSG*$^+(s)$ *has* $\|z\|_1 \leq \frac{(1+\varepsilon)\log ns}{\log\log ns}$ *for any constant* $\varepsilon > 0$.

Proof The RMG$^+(s)$ case has been proved in [1] and the RSG$^+(s)$ case can be proved in a similar way too. Here we provide a quick proof based on [1]'s proof to bypass the long algebra.

Let θ be the largest 1-norm of the columns in M_σ. According to the properties of a RSG$^+(s)$, the probability of $\theta > n - 1$ is 0 and $\Pr[\theta = n] \leq n \cdot (n-1)^{-(n-1)}$ is exponentially small. For any $k < n - 1$,

$$\begin{aligned}
& \Pr[\theta \geq k] \leq n \cdot \Pr[\text{a particular column has 1-norm at least } k] \\
\leq & n \cdot \binom{n-1}{k}\left(\frac{1}{n-1}\right)^k \leq 2(n-1) \cdot \binom{n-1}{k}\left(\frac{1}{n-1}\right)^k
\end{aligned}$$

Angluin and Chen [1] proved when $k = \frac{(1+\varepsilon)\log ns}{\log\log ns}$, $n \cdot \binom{n}{k}\left(\frac{1}{n}\right)^k = \frac{1}{s} \cdot o(1)$. Thus, in our case when $k = \frac{(1+\varepsilon)\log(n-1)s}{\log\log(n-1)s} \leq \frac{(1+\varepsilon)\log ns}{\log\log ns}$, we have $(n-1) \cdot$

$\binom{n-1}{k}\left(\frac{1}{n-1}\right)^k = \frac{1}{s} \cdot o(1)$ so that $\Pr[\theta \geq k] \leq \frac{1}{s} \cdot o(1)$. There are in total s matrices $\{M_\sigma \mid \sigma \in \Sigma\}$. Using a union bound we have $\|z\|_1 \leq \frac{(1+\varepsilon)\log ns}{\log\log ns}$ for all columns in all M_σ with probability $1 - o(1)$. ■

Appendix F Undirected connectivity

For strongly connected graphs such as RDG(s) and RG(s), the algorithm reconstructs the whole graph. If the target graph is not irreducible on the entire graph, e.g., RMG$^+$(s) and RSG$^+$(s), it recovers only the irreducible component of the graph because it relies on the convergence of the random walk and any vertex not in the irreducible component will have zero probability after convergence. We have no information for reconstructing the disconnected part. Here we prove that all random regular graphs in Table 1 have undirected connectivity on the entire graph with high probability. This positive fact could be helpful for future works on recovering the entire graph (perhaps with different techniques or given data).

Lemma 14 *With probability* $1 - o(1)$*, a random regular graph has no isolated component, i.e., the underlying undirected graph is connected, where* $s \geq 3$ *for RG(s) and* $s \geq 2$ *for the other models in Table 1.*

Proof The undirected connectivity is obvious for RMG$^\pm$(s), RSG$^\pm$(s), RDG(s) and RG(s) as they are all strongly connected on the entire graph with high probability. Since the in-regular models are transpose of the out-regular models and RSG$^+$(s) can be generated from RMG$^+$(s) using the two-stage procedure described in the proof of Lemma 3, it only remains to show the undirected connectivity of RMG$^+$(s).

As we can always find a small constant $\epsilon > 0$ such that $\log_n\left(\frac{1+\epsilon}{s-1}\log_2 n\right) < \frac{s-1}{2s-1}$, we pick a value c such that $\log_n\left(\frac{1+\epsilon}{s-1}\log_2 n\right) \leq c \leq \frac{s-1}{2s-1}$. Denoted by $p_{iso}(m)$ the probability of existence of an isolated component of size m in the graph. Without loss of generality, we assume $m \leq \frac{n}{2}$ because once the graph is undirected disconnected, there must be at least one isolated component of size $\leq \frac{n}{2}$. By union bound the probability of undirected disconnectivity is upper

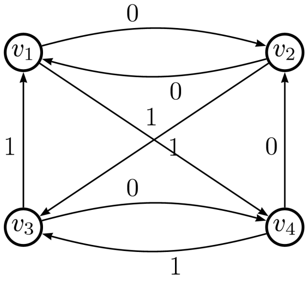

Figure 1: A 2-regular digraph with 4 vertices

bounded by the sum $\sum_{m=1}^{n/2} p_{iso}(m)$. For a particular integer $1 \le m \le \frac{n}{2}$,

$$\begin{aligned}
p_{iso}(m) &\le \binom{n}{m}\left(\frac{m}{n}\right)^{sm}\left(\frac{n-m}{n}\right)^{s(n-m)} \\
&\le \left(\frac{en}{m}\right)^{m}\left(\frac{m}{n}\right)^{sm}\left(\exp\left(-\frac{m}{n}\right)\right)^{s(n-m)} \\
&\le \left(\frac{en}{m}\right)^{m}\left(\frac{m}{n}\right)^{sm}\left(\exp\left(-\frac{m}{n}\right)\right)^{2(n-\frac{n}{2})} \\
&= e^{m}\left(\frac{m}{n}\right)^{(s-1)m}\left(\exp\left(-\frac{m}{n}\right)\right)^{n} \\
&= \left(\frac{m}{n}\right)^{(s-1)m}
\end{aligned}$$

When $m = 1$, it's easy to see $p_{iso}(1) \le n^{1-s}$.

When $2 \le m \le n^c$, $p_{iso}(m) \le \left(\frac{n^c}{n}\right)^{2(s-1)} = n^{2(c-1)(s-1)}$.

When $n^c \le m \le n/2$, $p_{iso}(m) \le \left(\frac{n/2}{n}\right)^{(s-1)n^c} = 2^{(1-s)n^c}$.

Hence, $\sum_{m=1}^{n/2} p_{iso}(m) \le n^{1-s} + (n^c - 1)n^{2(c-1)(s-1)} + (n/2 - n^c)2^{(1-s)n^c} = O(n^{1-s}) + O(n^{-\epsilon})$ for $\log_n\left(\frac{1+\epsilon}{s-1}\log_2 n\right) \le c \le \frac{s-1}{2s-1}$ and we have the undirected connectivity. ■

Appendix G A toy example

Suppose we consider the 2-regular digraph in Figure 1 whose transition matrices are

$$M_0 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \text{ and } M_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

For any regular digraph, the stationary distribution p_λ is always the uniform distribution. As $\log_s n = \log_2 4 = 2$, the coefficient matrix P_A is

$$P_A = \begin{pmatrix} p_{00} \\ p_{01} \\ p_{10} \\ p_{11} \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.75 & 0.25 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.25 & 0.25 \end{pmatrix}$$

Denote by $z = (M_0(1,1), M_0(2,1), M_0(3,1), M_0(4,1))^\top$ the the first column of matrix M_0. Let vector b be $(p_{000}(1), p_{010}(1), p_{100}(1), p_{110}(1))^\top = (0.5, 0, 0.5, 0)^\top$ as defined in the algorithm. The algorithm recovers z by solving the equation system $P_A z = b$, that is, solving

$$\begin{cases} 0.5M_0(1,1) + 0.5M_0(2,1) + 0M_0(3,1) + 0M_0(4,1) = 0.5 \\ 0M_0(1,1) + 0M_0(2,1) + 0.75M_0(3,1) + 0.25M_0(4,1) = 0 \\ 0M_0(1,1) + 0.5M_0(2,1) + 0M_0(3,1) + 0.5M_0(4,1) = 0.5 \\ 0.5M_0(1,1) + 0M_0(2,1) + 0.25M_0(3,1) + 0.25M_0(4,1) = 0 \end{cases}$$

Similarly the algorithm recovers all columns in M_0 and M_1 and reconstructs the target graph. Note that in the statistical query model the above equation system is perturbed but we showed the algorithm is robust to statistical query noise.

Appendix H Experimental results

H.1 Estimate of $|||P_A^\dagger|||_\infty$

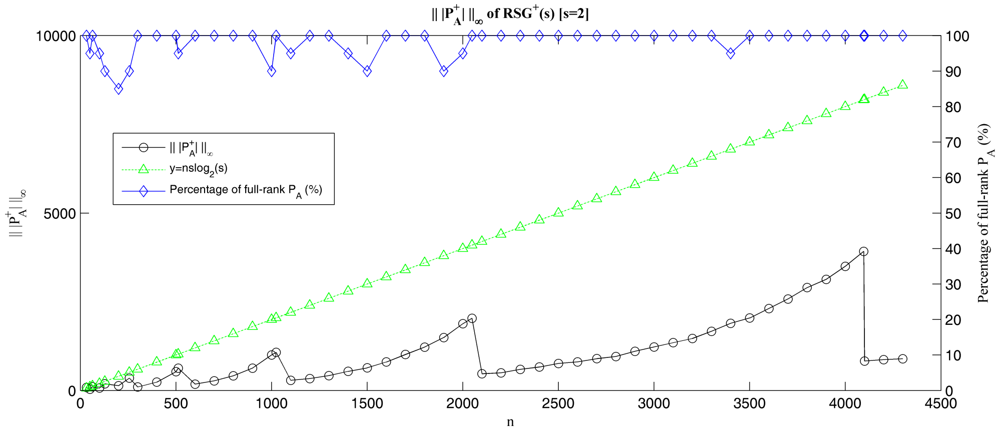

Figure 2: $|||P_A^\dagger|||_\infty$ of RSG$^+(s)$, versus n with fixed $s = 2$

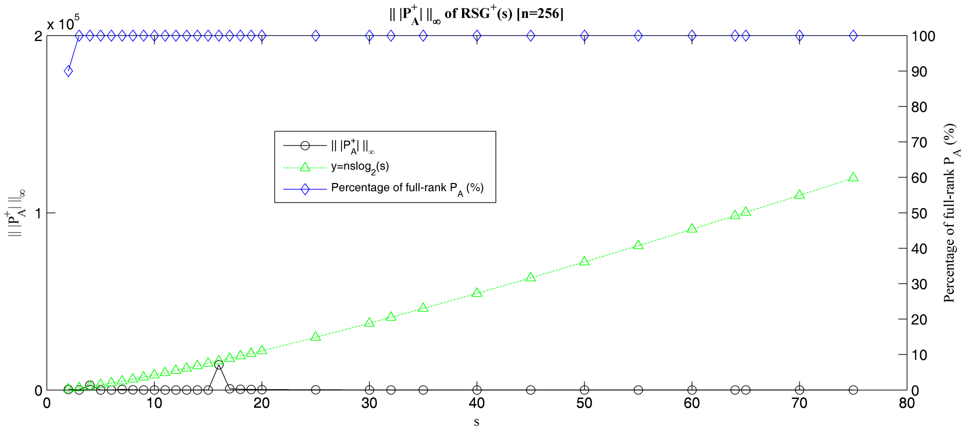

Figure 3: $|||P_A^\dagger|||_\infty$ of RSG$^+(s)$, versus s with fixed $n = 256$

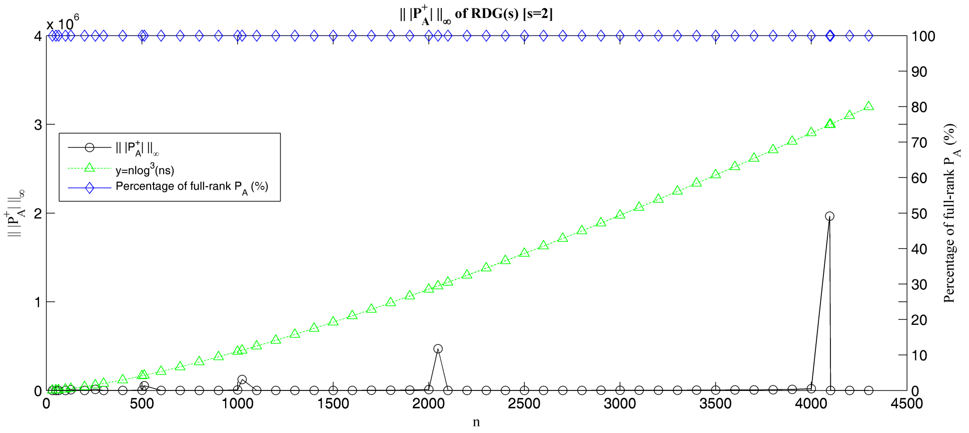

Figure 4: $\||P_A^\dagger|\|_\infty$ of RDG(s), versus n with fixed $s = 2$

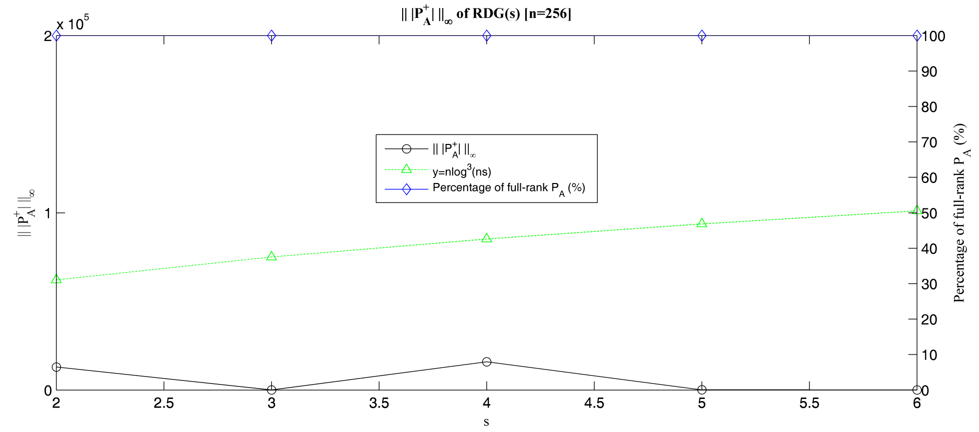

Figure 5: $\||P_A^\dagger|\|_\infty$ of RDG(s), versus s with fixed $n = 256$

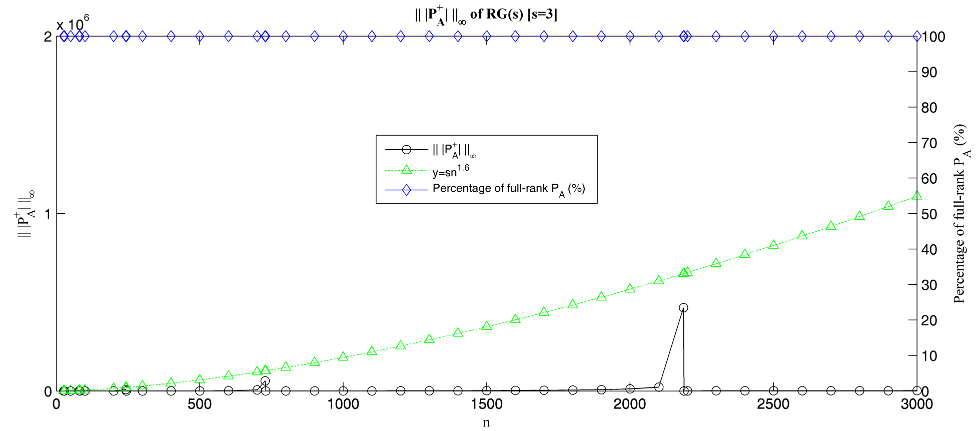

Figure 6: $\||P_A^\dagger|\|_\infty$ of RG(s), versus n with fixed $s = 3$

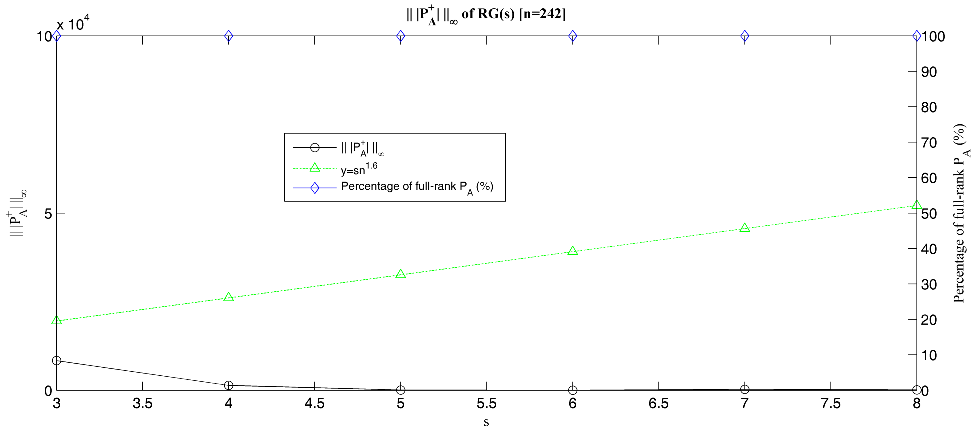

Figure 7: $\||P_A^\dagger|\|_\infty$ of RG(s), versus s with fixed $n = 242$

H.2 Maximum absolute error

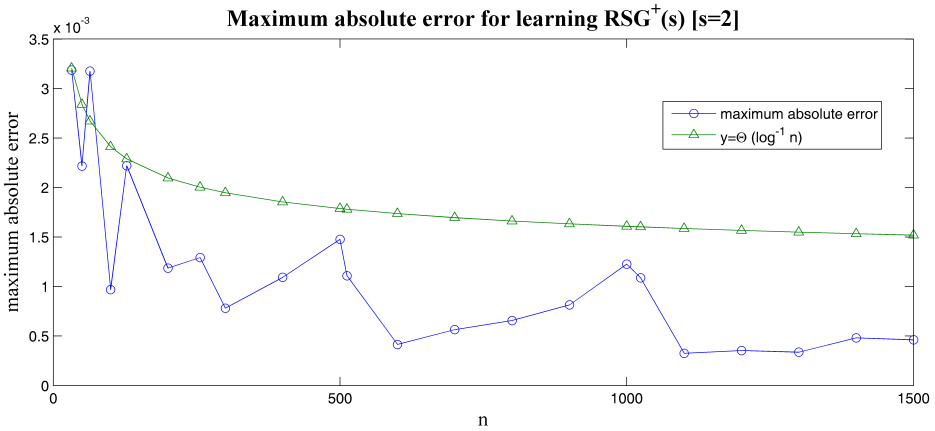

Figure 8: Maximum absolute error for learning RSG$^+(s)$, versus n with fixed $s = 2$

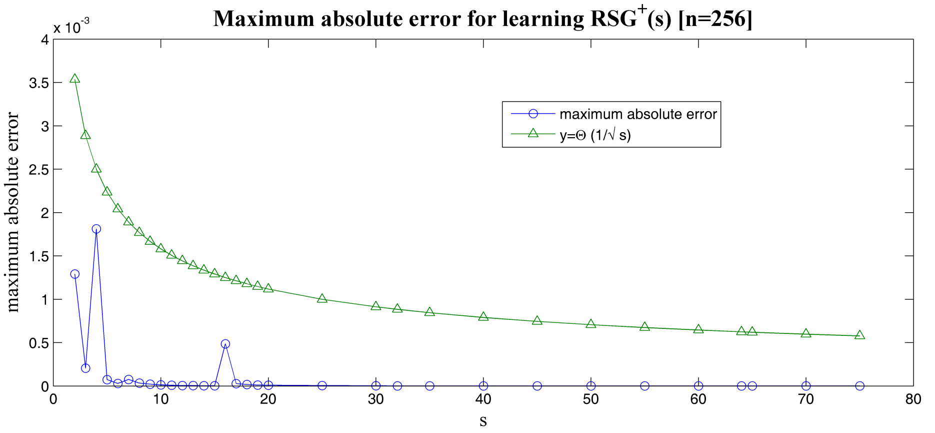

Figure 9: Maximum absolute error for learning RSG$^+(s)$, versus s with fixed $n = 256$

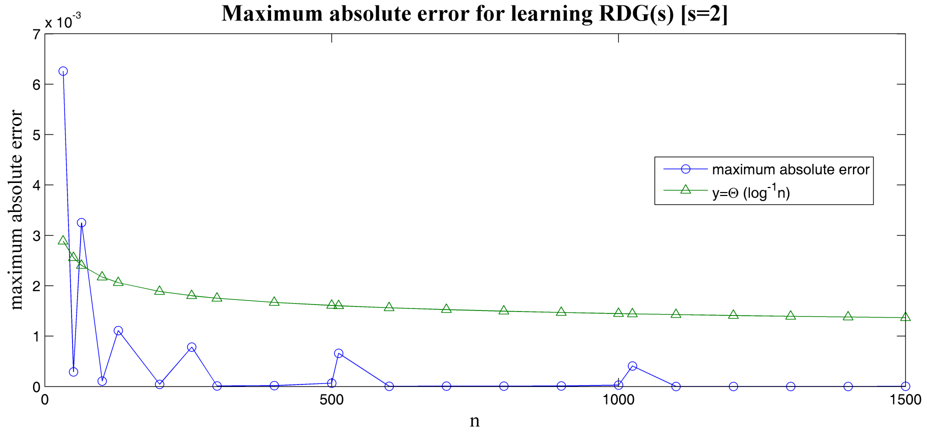

Figure 10: Maximum absolute error for learning RDG(s), versus n with fixed $s = 2$

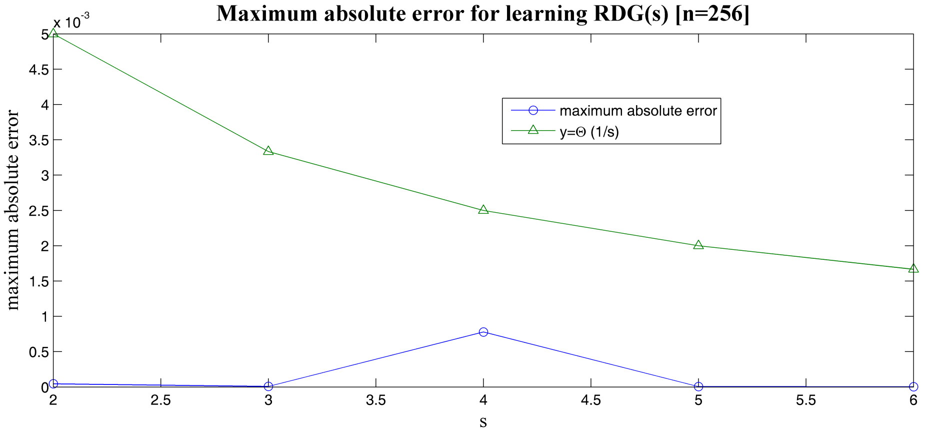

Figure 11: Maximum absolute error for learning RDG(s), versus s with fixed $n = 256$

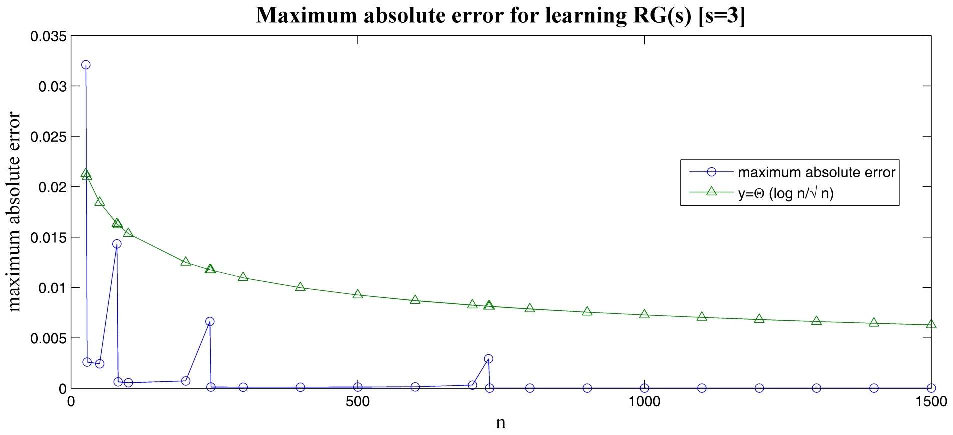

Figure 12: Maximum absolute error for learning RG(s), versus n with fixed $s = 3$

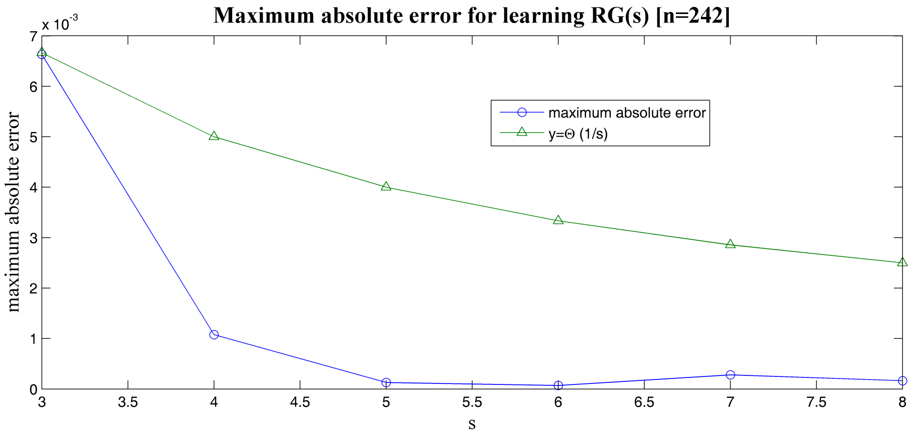

Figure 13: Maximum absolute error for learning RG(s), versus s with fixed $n = 242$

References

[1] D. Angluin and D. Chen. Learning a random DFA from uniform strings and state information. In *Proceedings of the 26th International Conference on Algorithmic Learning Theory*, 2015.

[2] E. A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Mathematics*, 10(2):217 – 223, 1974.

[3] M. A. Bender, A. Fernández, D. Ron, A. Sahai, and S. Vadhan. The power of a pebble: Exploring and mapping directed graphs. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 269–278. ACM, 1998.

[4] M. A. Bender and D. K. Slonim. The power of team exploration: Two robots can learn unlabeled directed graphs. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 75–85. IEEE, 1994.

[5] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311 – 316, 1980.

[6] V. Bourassa and F. Holt. Swan: Small-world wide area networks. In *Proceeding of International Conference on Advances in Infrastructures (SSGRR 2003w), LAquila, Italy*, 2003.

[7] M. Bui, T. Bernard, D. Sohier, and A. Bui. Random walks in distributed computing: A survey. In T. Bhme, V. Larios Rosillo, H. Unger, and H. Unger, editors, *Innovative Internet Community Systems*, volume 3473 of *Lecture Notes in Computer Science*, pages 1–14. Springer Berlin Heidelberg, 2006.

[8] M. B. D. Canfield, E. Rodney. Asymptotic enumeration of dense 0-1 matrices with equal row sums and equal column sums. *The Electronic Journal of Combinatorics [electronic only]*, 12(1):null, 2005.

[9] F. Chung. Laplacians and the Cheeger inequality for directed graphs. *Annals of Combinatorics*, 9:1–19, 2005.

[10] C. Cooper and A. Frieze. The cover time of random regular graphs. *SIAM Journal on Discrete Mathematics*, 18(4):728–740, 2005.

[11] C. Cooper, A. Frieze, and B. Reed. Random regular graphs of non-constant degree: connectivity and hamiltonicity. *Combinatorics, Probability & Computing*, 11(03):249–261, 2002.

[12] T. I. Fenner and A. M. Frieze. On the connectivity of random m-orientable graphs and digraphs. *Combinatorica*, 2(4):347–359, 1982.

[13] M. Fiedler. Bounds for eigenvalues of doubly stochastic matrices. *Linear Algebra and Its Applications*, 5(3):299–310, 1972.

[14] P. Fraigniaud, D. Ilcinkas, G. Peer, A. Pelc, and D. Peleg. Graph exploration by a finite automaton. pages 451–462, 2004.

[15] A. Grusho. Limit distributions of certain characteristics of random automaton graphs. *Mathematical notes of the Academy of Sciences of the USSR*, 14(1):633–637, 1973.

[16] M. Hildebrand. Random walks on random regular simple graphs. In *IMA Preprint Series*, 1994.

[17] M. Kearns. Efficient noise-tolerant learning from statistical queries. *J. ACM*, 45(6):983–1006, Nov. 1998.

[18] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald. Random regular graphs of high degree. *Random Structures & Algorithms*, 18(4):346–363, 2001.

[19] B. D. McKay. Asymptotics for 0-1 matrices with prescribed line sums. *Enumeration and Design,(Academic Press, 1984)*, pages 225–238, 1984.

[20] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica*, 11(4):369–382, 1991.

[21] G. Pandurangan, P. Raghavan, and E. Upfal. Building low-diameter peer-to-peer networks. *Selected Areas in Communications, IEEE Journal on*, 21(6):995–1002, 2003.

[22] O. Reingold. Undirected st-connectivity in log-space. In *Proceedings of the Thirty-seventh Annual ACM Symposium on Theory of Computing*, STOC '05, 2005.

[23] H.-A. Rollik. Automaten in planaren graphen. In *Theoretical Computer Science 4th GI Conference*, pages 266–275. Springer, 1979.

[24] P. Sarkar and A. Moore. Random walks in social networks and their applications: A survey. In C. C. Aggarwal, editor, *Social Network Data Analytics*, pages 43–77. Springer US, 2011.

[25] C. E. Shannon. Presentation of a maze-solving machine. In *8th Conf. of the Josiah Macy Jr. Found.(Cybernetics)*, pages 173–180, 1951.

[26] G. Thierrin. Permutation automata. *Theory of Computing Systems*, 2(1):83–90, 1968.

[27] B. Trakhtenbrot and I. Barzdin. *Finite Automata; Behavior and Synthesis.* Fundamental Studies in Computer Science, V. 1. North-Holland Publishing Company; New York: American Elsevier, 1973.

[28] N. C. Wormald. The asymptotic distribution of short cycles in random regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):168 – 182, 1981.

[29] N. C. Wormald. Models of random regular graphs. In J. D. Lamb and D. A. Preece, editors, *Surveys in Combinatorics, 1999*, pages 239–298. Cambridge University Press, 1999. Cambridge Books Online.