Generating Random Networks Without Short Cycles

Mohsen Bayati
Graduate School of Business, Stanford University, Stanford, CA 94305, bayati@stanford.edu

Andrea Montanari
Departments of Electrical Engineering and Statistics, Stanford, CA 94305, montanar@stanford.edu

Amin Saberi
Departments of Management Science and Engineering and Institute for Computational and Mathematical Engineering, Stanford, CA 94305, saberi@stanford.edu

Random graph generation is an important tool for studying large complex networks and has been extensively used by biologists, computer scientists, economists, electrical engineers, and mathematical sociologists. Despite abundance of models for random networks, unfortunately, the problem of efficiently generating random networks with application-driven constraints is poorly understood.

In order to advance state-of-the-art in this complex topic, we focus on a stylized family of graphs, random graphs without short cycles and develop new methodology. In particular, we construct a simple and efficient algorithm, RandGraph, for randomly generating simple graphs without small cycles. For any constant $k$, when $m = O(n^{1+1/[2k(k+3)]})$, RandGraph generates an asymptotically uniform random graph with $n$ vertices, $m$ edges, and no cycle of length $k$ or smaller using $O(n^2m)$ operations. To the best of our knowledge, this is the first polynomial-time algorithm for the problem. RandGraph works by sequentially adding $m$ edges to an empty graph with $n$ vertices. Recently, such sequential algorithms have been successful for random sampling problems. Our contribution to this line of research is development of a new methodology for characterizing an inherent bias in sequential algorithms and obtaining tighter concentration bounds for output distribution of these algorithms.

Key words: Network models, Poisson approximation, Sampling, Stochastic networks

1. Introduction

Recently, a common objective in many application areas has been extracting information from data sets that contain a network structure. Examples of such data are the Internet, social networks, biological networks, or healthcare networks such as network of physician referrals. In the last example, consider the question “how is the network of physician referrals formed?” Answering
this question could allow policy makers to influence the formation of the network with the objective of improving quality of care. This could be achieved by rewarding referrals to higher quality physicians and penalizing referrals to lower performing physicians. Unfortunately, empirical analysis of such network related questions is challenging since in most cases researchers have access to a single network or a few snapshots of it over time. Specifically, this “small \( n \)” problem renders the estimation part of any parametric network formation model unreliable (Chandrasekhar 2015).

A popular approach in statistical data analysis, when facing small number of observations, is bootstrap (Efron 1979) which increases the number of observations by creating random resamples of the original data. However, creating random copies of the networks is computationally expensive. For example, if we are interested in creating a random copy of the physician referral network while keeping the number of neighbors (degree) of each node fixed, the problem becomes NP hard in general (Wormald 1999). The property of fixing the number of neighbors is relevant when the researcher aims to control for the fact that certain physicians are more social than the other ones. Similarly, one could be interested in creating random copies of a network when certain substructures should be preserved or avoided. This problem in general is unsolved from a theoretical point of view except for few examples where efficient algorithms are proposed (Wormald 1999). Therefore, practitioners use non-rigorous heuristic models of random networks which may lead to incorrect (biased) estimates, see (Milo et al. 2002) for such a heuristic.

The objective of this paper is to advance state-of-the-art in this line of research by proposing a new algorithmic and analysis technique. We present the approach for a stylized subclass of problems, namely generating random graphs without short cycles and leave the extensions to other substructures for future research. While our emphasis in this paper is on advancing the methodology, and the family of graphs without short cycles is selected as an example of open problems in this area, we note that randomly generating graphs from this family has practical implications in information theory. Such graphs are used in designing low density parity check (LDPC) codes that can achieve Shannon capacity for transmitting messages in a noisy environment.
1.1. Contributions

We present a simple and efficient algorithm, RandGraph, for randomly generating simple graphs without short cycles. For any constant $k$, $\alpha \leq 1/[2k(k+3)]$ and $m = O(n^{1+\alpha})$, RandGraph generates an asymptotically uniform random graph with $n$ vertices, $m$ edges, and no cycle of length $k$ or smaller. RandGraph uses $O(n^2m)$ operations in expectation. To the best of our knowledge, this is the first polynomial-time algorithm for the problem.

RandGraph starts with an empty graph and sequentially adds $m$ edges between pairs of non-adjacent vertices. In every step, two distinct vertices $i$, $j$ with distance at least $k$ are selected with probability $p_{ij}$, and an edge $(ij)$ is added to the graph. The most crucial step, computing $p_{ij}$, is obtained by finding a sharp estimate for the number of valid extensions of the partially constructed graph, $G_t$, that contain $(ij)$ and have no cycle of length at most $k$. This estimation is done by computing the expected number of small cycles produced if the rest of the edges are added uniformly at random, using a Poisson approximation.

Our analysis of RandGraph relies on three key approximation methods. First we approximate random graphs that have $m$ edges and $n$ vertices with Erd"{o}s-Rényi (ER) graphs where each edge appears independently with probability $m/(\binom{n}{2})$. The second approximation uses Janson’s inequality [Janson 1990] for estimating the probability that random ER graphs have no cycle of length at most $k$. These two approximations provide us with an estimate for the uniform distribution on the family of graphs without cycles of length at most $k$. Finally, in the third step we approximate $G_t$ with ER graphs with edge density $t/m$, in order to estimate the output distribution of RandGraph, and to show that it is asymptotically equal to the uniform distribution. We emphasize that these approximations are easy when $m = O(n)$, and our main contribution is to show that they are sharp.

We note that using the Poisson approximation method in §6.2 of [Janson et al. 2000] one can estimate this probability with an additive error that converges to 0 with a rate that is inversely polynomial in $n$. However, we need a stronger approximation in this paper since we need a multiplicative error that converges to 1. This would require the additive error to converge to zero faster than the probability of the event itself which is exponentially small in $n$ when $m = O(n^{1+\alpha})$. 
even when the number of edges is super-linear in $n$, namely when $m = O(n^{1+\alpha})$ for small values of $\alpha$.

1.2. Organization of the Paper

The rest of the paper is organized as follows. §2 discusses related research. Description of RandGraph and statements of the main results are presented in §3. §4 provides the main idea behind RandGraph followed by its analysis in §5. An efficient implementation of RandGraph is presented in §6 and an extension to bipartite graphs with given degrees is given in §7.

2. Related Literature

Random graph models have been used in a wide variety of research areas. For example they are used in determining the effect of having overweight friends in adolescent obesity (Valente et al. 2009), in studying social networks that result from uncoordinated random connections created by individuals (Jackson and Watts 2002), in modeling emergence of the world wide web as an endogenous phenomena (Papadimitriou 2001) with certain topological properties (Kleinberg 2000, Newman 2003), and in simulating networking protocols on the Internet topology (Tangmunarunkit et al. 2002, Faloutsos et al. 1999, Medina et al. 2000, Bu and Towsley 2002).

Random graphs are also used in information theory (Richardson and Urbanke 2008). In particular, it has been shown that LDPC codes can approach Shannon capacity, asymptotically for large size codes, when their associated graph representations (Tanner graphs) are selected uniformly at random from the set of bipartite graphs with given degree sequences (Amraoui et al. 2007, Chung et al. 2001, Luby et al. 1997). While these random graphs guarantee optimal performances asymptotically, in practice the maximum graph has between $10^3$ and $10^5$ nodes where it is shown that the existence of a small number of subgraphs spoil the code performances (Di et al. 2002, Richardson 2003, Koetter and Vontobel 2003). In the present paper, we will focus on a specific class of such subgraphs, namely short cycles, however we expect our approach to be applicable to other subgraphs as well. In addition, for the sake of simplicity, we present the relevant proofs only
for the problem of generating random graphs without short cycles (not necessarily bipartite nor
with prescribing the degrees). Then we will adapt the algorithm for generating random bipartite
graphs with given degree sequences that have no short cycles\footnote{A more detailed discussion of the application to LDPC codes (including implementation details) can be found in this conference paper \cite{Bayati2009}.}. Generalizing proofs to this case is
cumbersome but we expect them to be conceptually straightforward.

Random graph generation has also been studied extensively as an interesting theoretical problem
\cite{Wormald1999, Ioannides2006}. From a theoretical perspective, our work is related to the following
problem. Consider a graph property $P$ that is preserved by removal of any edge from the graph.
It is a standard problem in extremal graph theory to determine the largest $m = m(n)$ such that
there exists a graph with $n$ vertices and $m$ edges having property $P$. Lower bounds on $m(n)$ can
be obtained through the analysis of greedy algorithms. Such algorithms proceed at each stage by
choosing uniformly among edges whose inclusion would not destroy property $P$, adding them to
the graph, and repeating the procedure until no further edges can be added. The resulting graph is
a random maximal $P$-graph. The question of finding the number of edges of a random maximal $P$-

\footnote{A more detailed discussion of the application to LDPC codes (including implementation details) can be found in this conference paper \cite{Bayati2009}.}
of RandGraph, a variant of the $C_3$-free process when $k = 3$, asymptotically converges to uniform in total variation distance early in the process; when $m$ is of order $n^{1+1/[2k(k+3)]}$. For this range of $m$ our result is sharper than the one by Pontiveros et al. (2013), see Remark 1 in §3 for details. Applying our analysis technique to prove similar results for the $C_{\ell}$-free process is an interesting area of future research. However, for the more general problem of generating graphs with a given degree sequence that have large girth as well, discussed in §7 we expect the natural extension of the $C_{\ell}$-free process to have an exponential bias. This is supported by the result of Bayati et al. (2010) that showed, when the degree sequence is irregular, the process of adding edges uniformly at random in the configuration model, while avoiding creation of double-edges or self-loops, generates graphs with a distribution that is asymptotically equal to the uniform distribution multiplied by an exponentially large bias\(^3\). However, Bayati et al. (2010) did not consider the constraint of avoiding small cycles. Therefore, studying whether the natural extension of the $C_{\ell}$-free process to graphs with prescribed degrees has such a bias is another interesting area of future research.

This paper is also closely related to recent literature on designing sequential algorithms for counting and generating random graphs with given degrees (Chen et al. 2005, Blitzstein and Diaconis 2010, Steger and Wormald 1999, Steger and Wormald 1999, Kim and Vu 2007, Bayati et al. 2010, Blanchet 2009). In fact, the current paper builds on this line of research and develops two mainly new techniques; (1) for obtaining probabilities $p_{ij}$, instead of starting from a biased heuristic and correcting its output distribution by modifying each $p_{ij}$ using algebraic calculations, we use Poisson approximation to obtain correct probabilities directly; and (2) for the analysis, we use graph approximation methods and Janson’s inequality to control the accumulated error from the sequential approximation of the probabilities $p_{ij}$. This leads to a tighter bound than the concentration results used in (Kim and Vu 2007, Bayati et al. 2010).

\(^3\)For regular graphs Steger and Wormald 1999, Kim and Vu 2007, Bayati et al. 2010 provide a positive result; the output distribution becomes asymptotically uniform when the degrees of are order $\sqrt{n}$.\)
3. Algorithm RandGraph and Main Result

In this section we start by introducing some notation and then present our algorithm (RandGraph) followed by the main theorem on its asymptotic performance.

The girth of a graph $G$ is defined to be the length of its shortest cycle. Let $\mathcal{G}_{n,m}$ denote the set of all simple graphs with $m$ edges over $n$ vertices and let $\mathcal{G}_{n,m,k}$ be the subset of graphs in $\mathcal{G}_{n,m}$ with girth greater than $k$. Throughout the paper $k$ is a constant and is independent of $n$ and $m$.

For any positive integer $s$, the set of integers $1, 2, \ldots, s$ is denoted by $[s]$. The complete graph with vertex set $[n]$ is denoted by $K_n$. For a graph $G$ with $n$ vertices we label its vertices by integers in $[n]$ and an edge that connects node $i$ to node $j$, when $i, j \in [n]$, is denoted by $(ij)$. All graphs considered in this paper are undirected which means $(ij)$ and $(ji)$ refer to the same edge.

RandGraph starts with an empty graph $G_0$ on $n$ vertices and at each step $t$, $t \in \{0, 1, \ldots, m-1\}$, an edge $(ij)$ is added to $G_t$ from $Q(G_t)$, the set of edges that their addition to $G_t$ does not create a cycle of length at most $k$. Then $G_{t+1}$ will be defined to be $G_t \cup (ij)$. The main technical step in RandGraph is that the edge $(ij)$ is selected randomly from $Q(G_t)$, according to a carefully constructed probability distribution that is denoted by $p(ij|G_t)$ and is given by

$$p(ij|G_t) \equiv \frac{1}{Z(G_t)} e^{-E_k(G_t, ij)},$$

where $Z(G_t) \equiv \sum_{(ij) \in Q(G_t)} e^{-E_k(G_t, ij)}$ is a normalizing term, $E_k(G_t, ij) \equiv \sum_{r=3}^{k} \sum_{\ell=0}^{r-2} N_{r, \ell}^{G_t, ij} q_{\ell}^{-1-\ell}$, $q_{\ell} \equiv \frac{m-t}{\binom{\ell}{2}-t}$, and $N_{r, \ell}^{G_t, ij}$ is the number of simple cycles (cycles that do not repeat a vertex) in $K_n$ that have length $r$, include $(ij)$, and include exactly $\ell$ edges of $G_t$. We will provide the intuition behind this complex-looking formula in §4. In addition, in §6 we will provide an efficient way of calculating $p(ij|G_t)$ using sparse matrix multiplication. Throughout the paper, to simplify the notation, in mathematical formula we will refer to RandGraph by the short notation RG.

By construction if RandGraph outputs a graph $G$, then $G$ is a member of $\mathcal{G}_{n,m,k}$. If RandGraph outputs FAIL the algorithm will be repeated till it produces a graph. We will show later that the probability of FAIL output vanishes asymptotically. Let $\mathbb{P}_{RG}(G)$ be the probability that RandGraph
Algorithm 1 RandGraph.

Input: $n$, $m$, $k$

Output: An element of $G_{n,m,k}$ or FAIL

set $G_0$ to be a graph over vertex set $[n]$ and with no edges

for each $t$ in $\{0, \ldots, m-1\}$ do

if $|Q(G_t)| = 0$ then

stop and return FAIL

else

sample an edge $(ij)$ with probability $p(ij|G_t)$, defined by Eq. (1) below

set $G_{t+1} = G_t \cup (ij)$

end if

end for

if the algorithm does not FAIL before $t = m−1$ then

return $G_m$

end if

does not FAIL and returns a graph $G$. Let also $P_U$ be the uniform probability on the set $G_{n,m,k}$; that is $P_U(G) = 1/|G_{n,m,k}|$. Our goal is to show that $P_{RG}(G)$ and $P_U(G)$ are very close in total variation distance. The total variation distance between two probability measures $P$ and $Q$ on a set $X$ is defined by $d_{TV}(P, Q) \equiv \sup \left\{ |P(A) − Q(A)| : A \subset X \right\}$. Now, we are ready to state the main result of the paper. Its proof is provided in §5.

Theorem 1. For $m = O(n^{1+\alpha})$, $m \geq n$, and a constant $k \geq 3$ such that $\alpha \leq 1/[2k(k+3)]$, the failure probability of RandGraph asymptotically vanishes and the graphs generated by RandGraph are approximately uniform. In other words,

$$\lim_{n \to \infty} P_{RG}(\text{FAIL}) = 0 \quad \text{and} \quad \lim_{n \to \infty} d_{TV}(P_{RG}, P_U) = 0.$$ 

Remark 1. Theorem 1 provides a sharper result than the ones provided by [Pontiveros et al., 2013]. For example, Lemma 2.1 by [Pontiveros et al., 2013] shows that $|Q(G_{TF,m})|$ is with high probability,
as $n \to \infty$, equal to $\mathbb{E}_0[|Q(G)|] \cdot [1 \pm 2e^{2m^2/n^3}n^{1/4}(\log n)^3]$ where $G_{TF,m}$ is the graph obtained in step $m$ of the triangle-free ($C_3$-free) process and $G$ is a random graph from $\mathbb{G}_{n,m}$. On the other hand, from Theorem 1 we can see that, as $n \to \infty$, with high probability $|Q(G_m)| = \mathbb{E}_0[|Q(G)|] \cdot [1 \pm o(1)]$.

To see this let $X$ be the random variable $|Q(G_m)|$ under $\mathbb{P}_{RG}$ and let $Y$ be the random variable $|Q(G)|$ under $\mathbb{P}_U$. Then using Proposition 4.7 of (Levin et al. 2008), there exist a coupling $(X,Y)$ such that $d_{TV}(\mathbb{P}_{RG},\mathbb{P}_U) = \mathbb{P}\{X \neq Y\}$. Using this fact and Theorem 1 with probability 1 as $n \to \infty$, we have $X = Y = \mathbb{E}_0[Y] \cdot [1 \pm o(1)]$ since $Y$ is very close to its expectation with an error that grows slower than $\mathbb{E}_0[Y]$. Pontiveros et al. (2013) provide similar asymptotic approximations for other graph parameters than $|Q(G_{TF,m})|$ as well. The same argument as above can be applied to show that our variants, when $m = O(n^{1+1/[2k(k+3)]})$, are sharper.

The next result shows a run-time guarantee for RandGraph and is proved in §6.

**Theorem 2.** Let $n$, $m$, and $k$ satisfy the conditions of Theorem 1. There exist an implementation of RandGraph that uses asymptotically $O(n^2m)$ operations in expectation.

### 4. The Intuition Behind RandGraph

In order to understand RandGraph, and in particular the calculations for $[p(ij|G_t)]$, it is instructive to examine the execution tree $T$ of a simpler version of RandGraph that sequentially adds $m$ random edges to an empty graph on $n$ vertices to obtain an element of $\mathbb{G}_{n,m}$ (without any attention to whether any short cycles are generated). Consider a rooted $m$-level tree where the root (the vertex in level zero) corresponds to the empty graph at the beginning of this sequential algorithm and level $t$ vertices correspond to all pairs $(G_t, \pi_t)$ where $G_t$ is a partial graph that can be constructed after $t$ steps, and $\pi_t$ is an ordering of its $r$ edges. There is a link (edge) in $T$ between a partial graph $(G_t, \pi_t)$ from level $r$ to a partial graph $(G_{t+1}, \pi_{t+1})$ from level $t+1$ if $G_t \subset G_{t+1}$ and the first $t$ edges of $\pi_t$ and $\pi_{t+1}$ are equal. Any path from the root to a leaf at level $m$ of $T$ corresponds to one possible way of sequentially generating a random graph in $\mathbb{G}_{n,m}$.

Let us denote those partial graphs $G_t$ that have girth greater than $k$ by “valid” graphs. Our goal is to reach a valid leaf in $T$, uniformly at random, by starting from the root and going down
the tree. A naïve approach could be repeating the above sequential algorithm as long as its output in step $m$ is not a valid leaf of $T$. However, when $m = O(n^{1+\alpha})$, the fraction of valid leaves is of order $\exp(-n^\alpha)$ (see §5 for details). Therefore, this naïve approach has an exponentially small chance of success; i.e., to produce a graph with girth larger than $k$. We should emphasize that the focus of this paper is not the simple regime of $m = O(n)$ since when $m = O(n)$, valid leaves form a constant fraction of leaves of $T$ which would make the naïve approach a desirable algorithm and using RandGraph would not be necessary.

To fix the problem with the naïve approach when $m = O(n^{1+\alpha})$, RandGraph is designed to approximate a general strategy for uniformly randomly generating valid leaves of $T$ (Sinclair 1993): at any step $t$, choose $G_{t+1} = G_t \cup (ij)$ with probability proportional to the number of valid leaves of $T$ among descendant of $(G_{t+1}, \pi_{t+1})$. Denote this probability by $p_{\text{true}}(G_{t+1}, \pi_{t+1})$. The main challenge for implementing this strategy is calculating $p_{\text{true}}(G_{t+1}, \pi_{t+1})$. In RandGraph we will approximate $p_{\text{true}}(G_{t+1}, \pi_{t+1})$ with $p(G_{t+1}, \pi_{t+1})$ as follows. Let $n_k(G_{t+1}, \pi_{t+1})$ denote the number of cycles of length at most $k$ in a leaf chosen uniformly at random among descendants of $(G_{t+1}, \pi_{t+1})$ in $T$. Note that $p_{\text{true}}(G_{t+1}, \pi_{t+1})$ is by definition is equal to $\mathbb{P}\{n_k(G_{t+1}, \pi_{t+1}) = 0\}$. Using Poisson approximation, see (Alon and Spencer 1992) for details, we can show that the distribution of $n_k(G_{t+1}, \pi_{t+1})$ is approximately Poisson. In particular,

$$\mathbb{P}\{n_k(G_{t+1}, \pi_{t+1}) = 0\} \approx \exp(-\mathbb{E}[n_k(G_{t+1}, \pi_{t+1})]) .$$

(2)

Therefore, our approximation $p(G_{t+1}, \pi_{t+1})$ will be chosen to be proportional to the right hand side of Eq. (2). This is the main intuition behind Eq. (1) and will be formally explained in §5.

A crucial step in the analysis of RandGraph, provided in §5 is to control the accumulated error

$$\prod_{t=0}^{m-1} \left[ \frac{p(G_{t+1}, \pi_{t+1})}{p_{\text{true}}(G_{t+1}, \pi_{t+1})} \right] ,$$

Prior work (Kim and Vu 2007, Bayati et al. 2010) used sharp concentration inequalities to find a separate upper bound, for each $r$, on the error term $[p(G_{t+1}, \pi_{t+1})/p_{\text{true}}(G_{t+1}, \pi_{t+1})]$. Instead, in this paper we simplify the final product $\prod_{t=0}^{m-1} [p(G_{t+1}, \pi_{t+1})/p_{\text{true}}(G_{t+1}, \pi_{t+1})]$ and will approximate it directly which leads to a tighter bound.
5. Analysis of RandGraph and Proof of Theorem [1]

The aim of this § is to prove Theorem [1]. The most crucial part of the proof is to show that probability of generating a graph $G$ by RandGraph, $P_{RG}(G)$ is asymptotically larger than the uniform probability over $G_{n,m,k}$ or $P_{U}(G)$. After this result is stated in Lemma [1] below, it is used to prove Theorem [1]. The rest of the section is divided into four subsections. In particular, §5.1 describes the main steps for proving Lemma [1] which rely on auxiliary Lemmas [2] and [3]. These auxiliary lemmas are stated in §5.1 and proved in §5.2 and §5.3 respectively. Throughout this section we will introduce a large number of new notations. For convenience, we have repeated all notations, required for understanding this analysis, with their definition in Table 1 in Appendix A.

**Lemma 1.** For any constant $\tau$, let $G_{n,m,k}(\tau)$ be the subset of graphs $G$ in $G_{n,m,k}$ such that

$$P_{RG}(G) < (1 - \tau)P_{U}(G).$$

For every $n, m, k$ satisfying the conditions of Theorem [1] there exist a constant $\tau_{n,m,k}$ such that as $n$ grows to $\infty$ we have $\tau_{n,m,k} \to 0$ and

$$\lim_{n \to \infty} \frac{|G_{n,m,k}(\tau_{n,m,k})|}{|G_{n,m,k}|} = 0. \quad (3)$$

In other words, Lemma [1] shows that for all but $o(|G_{n,m,k}|)$ graphs $G$ in $G_{n,m,k}$ inequality $P_{RG}(G) \geq [1 - o(1)]P_{U}(G)$, holds where the term $o(1)$ goes to zero as $n$ goes to infinity uniformly in the graph $G$. Now we are ready to prove our main Theorem, assuming Lemma [1].

**Proof of Theorem [1]** From the definition of $d_{TV}(P_{RG}, P_{U})$, using triangle inequality, we obtain

$$d_{TV}(P_{RG}, P_{U}) \leq \sum_{G \in G_{n,m,k}} |P_{RG}(G) - P_{U}(G)|.$$  

Then, depending on whether $P_{RG}(G) \geq P_{U}(G)$ or $P_{RG}(G) < (1 - \tau_{n,m,k})P_{U}(G)$ we bound the term $|P_{RG}(G) - P_{U}(G)|$ differently. Let $B_{n,m,k} \subset G_{n,m,k}$ be the set of all graphs $G$ with $P_{RG}(G) \leq P_{U}(G)$ and let the subset $D_{n,m,k} \subset B_{n,m,k}$ to be those graphs $G$ in $B_{n,m,k}$ with $P_{RG}(G) < (1 -$
\( \tau_{n,m,k} \mathbb{P}_U(G) \). To simplify the notation, for the rest of the proof we drop the subscripts \( n, m, k \) from \( \tau_{n,m,k}, \mathbb{B}_{n,m,k}, \mathbb{D}_{n,m,k} \) and \( G_{n,m,k} \). Assuming Lemma 1 holds then

\[
\lim_{n \to \infty} \frac{|D|}{|G|} = 0, \tag{4}
\]

and for \( G \in \mathbb{B} \setminus \mathbb{D} \)

\[
|P_{RG}(G) - P_U(G)| = P_U(G) - P_{RG}(G) \leq \tau P_U(G). \tag{5}
\]

Therefore,

\[
\sum_{G \in G} |P_{RG}(G) - P_U(G)| = \sum_{G \in G} |P_{RG}(G) - P_U(G)| + 2 \sum_{G \in \mathbb{B}} |P_{RG}(G) - P_U(G)|
\]

\[
= \sum_{G \in G} |P_{RG}(G) - P_U(G)| + 2 \sum_{G \in \mathbb{B} \setminus \mathbb{D}} |P_{RG}(G) - P_U(G)| + 2 \sum_{G \in \mathbb{D}} |P_{RG}(G) - P_U(G)|
\]

\[
\leq \sum_{G \in G} P_{RG}(G) - \sum_{G \in G} P_U(G) + 2 \tau \sum_{G \in \mathbb{B} \setminus \mathbb{D}} P_U(G) + 2 \sum_{G \in \mathbb{D}} P_U(G)
\]

\[
\leq 1 - P_{RG}(\text{FAIL}) - 1 + 2 \tau + 2 \frac{|D|}{|G|}
\]

\[
\leq 2 \tau + 2 \frac{|D|}{|G|},
\]

where (a) uses Eq. (5) and \( P_{RG}(\text{FAIL}) \) is the probability of failure of RandGraph. Now taking \( n \to \infty \) and using Eq. (4) and the fact that \( \lim_{n \to \infty} \tau \to 0 \) the result follows \( \Box \)

Throughout the rest of this section our focus will be on proving Lemma 1.

5.1. Lower Bound For \( P_{RG}(G) \): Proof of Lemma 1

This proof contains four main steps.

Step 1 in Proof of Lemma 1: Approximating \( P_U \). Since \( P_U = 1/|G_{n,m,k}| \), we will find an asymptotic estimate for \( |G_{n,m,k}| \) using Janson’s inequality (Janson 1990). Janson’s inequality shows the number of cycles of constant length in \( G_{n,m} \) is approximately a Poisson random variable. The result is summarized in the following lemma that is proved in §5.2. But before stating the lemma we define \( C_r \) to be the set of all simple cycles of length \( r \) in \( K_n \) and introduce notation \( N \) for total number of edges in \( K_n \) which is equal to \((n)\).
Lemma 2. Let \( m = O(n^{1+\alpha}) \) with \( \alpha < 1/(2k - 1) \) for \( k \geq 3 \), and \( m \geq n \), then

\[
\lim_{n \to \infty} \frac{|G_{n,m,k}|}{\binom{n}{m} \exp \left[ -\sum_{r=3}^{k} |C_r| \left( \frac{m}{n} \right)^r \right]} = 1.
\]

Therefore, we have the following asymptotic estimate for \( P_U(G) \)

\[
\frac{P_U(G)}{\left\{ \binom{n}{m} \exp \left[ -\sum_{r=3}^{k} |C_r| \left( \frac{m}{n} \right)^r \right] \right\}^{-1} \to 1, \quad \text{as } n \to \infty. \tag{7}
\]

The remaining steps will provide necessary approximations and algebraic simplifications to find an asymptotic lower bound for \( P_{RG} \) which will be equal to the denominator term in Eq. (7).

Step 2 in Proof of Lemma 2: Using convexity and Jensen’s Inequality. Let us start by writing an expression for \( P_{RG}(G) \) when \( G \) is a fixed element of \( G_{n,m,k} \). Note that RandGraph sequentially adds edges to an empty graph to produce a graph with \( m \) edges. Hence for the fixed graph \( G \), there are \( m! \) permutations of the edges of \( G \) that can be generated by RandGraph and each permutation can have a different probability. Let \( \pi \) be any permutation of the edges of \( G \) (i.e. a one-to-one mapping from \( \{1, \ldots, m\} \) to the edges of \( G \)), and let \( G^\pi_t \) be the graph having \( [n] \) as vertex set and \( \{\pi(1), \ldots, \pi(t)\} \) as edge set. This is the partial graph that is generated after \( t \) steps of RandGraph conditioned on having \( \pi \) as output. Now we can write

\[
P_{RG}(G) = \sum_{\pi} \prod_{t=0}^{m-1} p(\pi(t+1)|G^\pi_t).
\]

Additionally, consider the uniform distribution on the set of all \( m! \) permutations \( \pi \). Then, \( \sum_{\pi} \) can be replaced by \( m! \mathbb{E}_\pi \) where \( \mathbb{E}_\pi \) is expectation with respect to a random permutation \( \pi \). Hence,

\[
P_{RG}(G) = m! \mathbb{E}_\pi \left\{ \prod_{t=0}^{m-1} p(\pi(t+1)|G^\pi_t) \right\}
\]

\[
= m! \mathbb{E}_\pi \exp \left\{ \sum_{t=0}^{m-1} \log p(\pi(t+1)|G^\pi_t) \right\}
\]

\[
\geq m! \exp \left\{ \sum_{t=0}^{m-1} \mathbb{E}_\pi \log p(\pi(t+1)|G^\pi_t) \right\}, \tag{8}
\]

where the inequality is by Jensen’s inequality for the convex function \( e^x \).
Next, applying the definition of \( p(\pi(t+1)|G_t) \) from Eq. (1) we get

\[
\mathbb{P}_{\text{RG}}(G) \geq m! \exp \left[ -\sum_{t=0}^{m-1} \mathbb{E}_\pi E_k(G_t^\pi, \pi(t+1)) - \sum_{t=0}^{m-1} \mathbb{E}_\pi \log Z(G_t^\pi) \right].
\] (9)

Now, we define \( F_r(G_t^\pi) \) to be the number of all forbidden pairs at step \( t \), pairs of nodes \( i \) and \( j \) such that adding \((ij)\) to \( G_t^\pi \) creates a cycle of length \( r \), and set \( Z_0(G_t^\pi) \equiv N - t - \sum_{r=3}^k F_r(G_t^\pi) \). To simplify the notation, we drop the reference to \( G \) and obtain

\[
\log Z = \log Z_0 + \log \left( \frac{Z}{Z_0} \right)
\]

\[
= \log \left( (N - t)(1 - \sum_{r=3}^k \frac{F_r}{N - t}) \right) + \log \left( \frac{Z}{Z_0} \right)
\]

\[
\leq \log(N - t) - \sum_{r=3}^k \frac{F_r}{N - t} + \log \left( \frac{Z}{Z_0} \right),
\] (10)

using inequality \( \log(1 - x) \leq -x \) for \( x \in (-\infty, 1] \) which is applicable since \( \sum_{r=3}^k F_r \leq N - t \). Combining Eqs. (9) and (10), we arrive at the following modified lower bound for \( \mathbb{P}_{\text{RG}}(G) \)

\[
\mathbb{P}_{\text{RG}}(G) \geq \frac{1}{\binom{N}{m}} \exp \left[ -\sum_{t=0}^{m-1} \mathbb{E}_\pi E_k(G_t^\pi, \pi(t+1)) + \frac{1}{N} \sum_{r=3}^k \sum_{t=0}^{m-1} \mathbb{E}_\pi F_r(G_t^\pi) - \sum_{t=0}^{m-1} \mathbb{E}_\pi \log \left( \frac{Z(G_t^\pi)}{Z_0(G_t^\pi)} \right) \right].
\] (11)

Note that \( S_1(G) \) and \( S_3(G) \) include the negative sign as well as the summation term.

The next step is the most important part of our effort in the journey to prove Lemma 1.

**Step 3 in Proof of Lemma 1:** Simplifying the exponent \( S_1(G) + S_2(G) + S_3(G) \). This step shows the main benefit of deferring the calculation of approximation errors for \( p(ij|G_t^\pi) \) to the final step. We will show that even though the terms \( S_i(G) \) for \( i = 1, 2, 3 \) can be large and dependent on \( G \), many terms in their combined sum cancel out and the resulting expression will be independent of \( G \). In particular, we will show that the only negative term, \( S_1 \), will completely cancel \( S_2 \) and all graph dependent parts of \( S_3 \). Throughout the rest, since \( G \) is fixed, we often drop the references to \( G \) in \( S_i : i = 1, 2, 3 \).

The main result of this step is summarized in the following lemma. First we define \( C_{r, \ell}(G) \) to be the set of all simple cycles of length \( r \), belonging to \( K_n \), that include exactly \( \ell \) edges of \( G \).
Lemma 3. Let \( m = O(n^{1+\alpha}) \) and constant \( k \geq 3 \) be such that \( \alpha \leq 1/[2k(k+3)] \) and \( m \geq n \). Then there exist a constant \( \tau_{n,m,k}^{(1)} \) such that \( \tau_{n,m,k}^{(1)} \to 0 \) as \( n \) goes to \( \infty \) and the following inequalities hold for all but a vanishing fraction of graphs \( G \) in \( \mathbb{G}_{n,m,k} \) (i.e., the subset of \( \mathbb{G}_{n,m,k} \) that does not satisfy the inequalities has a size of \( o(|\mathbb{G}_{n,m,k}|) \)).

(a) \( S_1(G) \geq \tau_{n,m,k}^{(1)} - \sum_{r=3}^{k} \sum_{\ell=1}^{r-1} |C_{r,\ell}(G)| \left( \frac{m}{N} \right)^{r-\ell} \int_{0}^{1} \theta^{\ell-1}(1-\theta)^{r-\ell} d\theta \).

(b) \( S_2(G) \geq \tau_{n,m,k}^{(1)} + \sum_{r=3}^{k} |C_{r,r-1}(G)| \frac{m}{N} \int_{0}^{1} \theta^{-1} d\theta \).

(c) \( S_3(G) \geq \tau_{n,m,k}^{(1)} + \sum_{r=3}^{k} \sum_{\ell=0}^{r-2} |C_{r,\ell}(G)| \left( \frac{m}{N} \right)^{r-\ell}(r-\ell) \int_{0}^{1} \theta^{\ell}(1-\theta)^{r-\ell-1} d\theta \).

We defer proof of Lemma 3 to §5.3.

Final Step in Proof of Lemma 1. Next we will show how the different terms (lower bounds for \( S_i \)'s) in Lemma 3 cancel each other. The main idea in relating the terms in the lower bounds is the following equation which is obtained using integration by parts for \( r-1 \geq \ell > 1 \)

\[
\ell \int_{0}^{1} \theta^{\ell-1}(1-\theta)^{r-\ell} d\theta = (r-\ell) \int_{0}^{1} \theta^{\ell}(1-\theta)^{r-\ell-1} d\theta. \tag{12}
\]

Using (12) we can see that, when adding the right hand sides of the three inequalities in Lemma 3 all terms in the lower bound for \( S_1 \) with \( 1 \leq \ell \leq r-2 \) are canceled with the corresponding terms in the lower bound for \( S_3 \). In addition, the \( \ell = r-1 \) term in the lower bound of \( S_1 \) is canceled with the lower bound of \( S_2 \). Therefore, the uncanceled terms are \( \ell = 0 \) terms from the lower bound of \( S_3 \) which we will see below to be asymptotically independent of \( G \). More formally, combining Eq. (11) and Lemma 3 for all graphs \( G \) in \( \mathbb{G}_{n,m,k} \), except a subset of size \( o(|\mathbb{G}_{n,m,k}|) \),

\[
\mathbb{P}_{RG}(G) \geq \left( \frac{N}{m} \right)^{-1} \exp \left[ S_1(G) + S_2(G) + S_3(G) + 3r_{n,m,k}^{(1)} \right]
\]

\[
\geq \left( \frac{N}{m} \right)^{-1} \exp \left[ \sum_{r=3}^{k} |C_{r,0}(G)| \left( \frac{m}{N} \right)^{r} \int_{0}^{1} (1-\theta)^{r-1} d\theta + 3r_{n,m,k}^{(1)} \right]
\]

\[
= \left( \frac{N}{m} \right)^{-1} \exp \left[ \sum_{r=3}^{k} |C_{r,0}(G)| \left( \frac{m}{N} \right)^{r} + 3r_{n,m,k}^{(1)} \right], \tag{13}
\]
with $\tau_{n,m,k}^{(1)}$ converging to 0 as $n$ goes to infinity.

Comparing (13) and the asymptotic expression for $\mathbb{P}_U(G)$ given by the denominator term in Eq. (7), obtained from Lemma 2, we see that the only difference in the exponent is the use of $|\mathcal{C}_{r,0}(G)|$ instead of $|\mathcal{C}_r|$ and the following lemma, proved in §B, provides the final piece.

**Lemma 4.** If $m = O(n^{1+\alpha})$ and $k$ is constant then $|\mathcal{C}_r \setminus \mathcal{C}_{r,0}(G)|/|\mathcal{C}_r| = O(n^{\alpha-1})$.

Using Lemma 4 we have

$$\sum_{r=3}^{k} |\mathcal{C}_{r,0}(G)| \left( \frac{m}{N} \right)^r \geq \sum_{r=3}^{k} |\mathcal{C}_r| \left[ 1 - O(n^{\alpha-1}) \right] \left( \frac{m}{N} \right)^r$$

$$\geq -O(n^{(k+1)\alpha-1}) + \sum_{r=3}^{k} |\mathcal{C}_r| \left( \frac{m}{N} \right)^r$$

$$= o(1) + \sum_{r=3}^{k} |\mathcal{C}_r| \left( \frac{m}{N} \right)^r,$$

where the second inequality uses $|\mathcal{C}_r| = O(n^r)$ and $m = O(n^{1+\alpha})$. Thus, there exist constants $\tau_{n,m,k}^{(2)}, \tau_{n,m,k}^{(3)}$, converging to 0 as $n$ goes to infinity, where

$$\mathbb{P}_{RG}(G) \geq \frac{1}{\binom{N}{m}} \exp \left[ \tau_{n,m,k}^{(2)} - \sum_{r=3}^{k} |\mathcal{C}_r| \left( \frac{m}{N} \right)^r \right] \geq (1 - \tau_{n,m,k}^{(3)}) \mathbb{P}_U(G),$$

for all graphs $G$ in $\mathcal{G}_{n,m,k}$, except a subset of size $o(|\mathcal{G}_{n,m,k}|)$. This finishes proof of Lemma 1. □

### 5.2. Approximating $|\mathcal{G}_{n,m,k}|$ and Proof of Lemma 2

Before delving into the details, we provide a high-level overview of the proof. The main idea is to look at the random graph model $\mathcal{G}_{n,m}$ and estimate the probability of the event of having a graph with girth larger than $k$ using Janson’s inequality. However, we will do all of this on an approximation to the random graph model $\mathcal{G}_{n,m}$, namely random graph model $\mathcal{G}_{n,p}$ where each edge on vertices of $[n]$ appears independently randomly with probability $p = m/N$. Any graph in $\mathcal{G}_{n,p}$ would have on average $m$ edges, making $\mathcal{G}_{n,p}$ a natural approximation to $\mathcal{G}_{n,m}$. We do this approximation in order to use Janson’s inequality which will be described next.
**Definition 1 (Janson’s Inequality).** Given a collection of “bad events” \{B_i : i \in I\}, we would like to estimate the probability \(P(\bigcap_{i \in I} B_i^{(c)})\) assuming that the events \(B_i^{(c)}, i \in I\) are “almost independent”. More formally, let \(\eta, \xi\) be real numbers such that \(\eta < 1\) and for all \(i, j\) in \(I\),

\[
P(B_i) \leq \eta \quad \text{and} \quad \sum_{B_j \sim B_i} P(B_i \cap B_j) = \xi.
\]

Here \(B_i \sim B_j\) means that \(B_i, B_j\) are dependent. Then Janson’s inequality is

\[
\prod_{i \in I} P(B_i^{(c)}) \leq P(\bigcap_{i \in I} B_i^{(c)}) \leq \prod_{i \in I} P(B_i^{(c)}) \exp \left( \frac{\xi}{2(1 - \eta)} \right). \tag{14}
\]

In particular, for \(\xi = o(1)\) we have \(P(\bigcap_{i \in I} B_i^{(c)}) = (1 + o(1)) \prod_{i \in I} P(B_i^{(c)})\).

Let us denote the probability with respect to the randomness in \(G_{n,p}\) and \(G_{n,m}\) by \(P_{n,p}\) and \(P_{n,m}\) respectively. Let \(A_k\) be the event that a random graph, selected from \(G(n, p)\) or \(G(m, n)\), has girth greater than \(k\). Our next step is to calculate \(P_{n,p}(A_k)\).

**5.2.1. Approximating \(P_{n,p}(A_k)\) via Janson Inequality.** By definition, \(P_{n,p}(A_k)\) is the probability that a random graph \(G\) in \(G_{n,p}\) has no cycle of length at most \(k\). For every cycle \(i\) of length at most \(k\) on vertices of \([n]\) we consider a bad event \(B_i\) that is the event that a random graph \(G\) from \(G_{n,p}\) contains cycle \(i\). In particular, \(I = \bigcup_{r=3}^{k} C_r\). It is not difficult to see that \(P(B_i) = O(p^3)\) and \(\xi = O(\sum_{r=3}^{k} n^{2r-2} p^{2r-1})\). And since \(p = O(n^{\alpha-1})\) then using Janson’s inequality (14),

\[
\prod_{i \in C} P(B_i^{(c)}) \leq P_{n,p}(A_k) \leq e^{O(n^{(2k-1)\alpha-1})} \prod_{i \in I} P(B_i^{(c)})
\]

which gives the following for \(\alpha < 1/(2k-1)\):

\[
P_{n,p}(A_k) = e^{o(1)} \prod_{i \in C} P(B_i^{(c)})
\]

\[
= e^{o(1)} \prod_{i \in C} (1 - p^{\text{length}(i)})
\]

\[
= \exp \left[ o(1) + \sum_{r=3}^{k} |C_r| \log(1 - p^r) \right]
\]

\[
= \exp \left[ o(1) - \sum_{r=3}^{k} |C_r| p^r \right]. \tag{15}
\]
The last equality uses $|C_r|^p 2^r = O(n^r n^{2r\alpha - 2r}) = o(1)$ since $\alpha < 1/2$.

Finally we are ready to prove Lemma 2 which is achieved by approximating $P_{n,m}(A_k)$ with $P_{n,p}(A_k)$. Before finalizing the proof, let us state the following result on monotone properties of random models $G_{n,p}$ and $G_{n,m}$. However, we only state it for the specific event $A_k$ but it is correct for any event with the following property; if $G \in A_k$ then any graph $G'$, obtained from $G$ via removal of an edge, is also contained in $A_k$. Such events are known as monotone graph properties.

**Proposition 1** (Janson-Luczak-Rucinski, Lemma 1.10 in [Janson et al. 2000]). For $0 \leq p \leq p' \leq 1$ and $0 \leq m \leq m' \leq N$ we have $P_{n,p}(A_k) \geq P_{n,p'}(A_k)$ and $P_{n,m}(A_k) \geq P_{n,m'}(A_k)$.

**Proof of Lemma 2.** First we note that for any $0 < p < 1$, the random graph model $G(n,p)$ is equivalent to the random graph model $G(n,m)$ when graphs are conditioned to have exactly $m$ edges. Thus for a random graph $G$ we have

$$P_{n,p}(A_k) = P_{n,p}(A_k \cap \{m(G) \geq m\}) + P_{n,p}(A_k \cap \{m(G) < m\})$$

$$\leq \sum_{\ell=m}^N P_{n,p}(A_k | m(G) = \ell) P_{n,p}(m(G) = \ell) + P_{n,p}(m(G) < m)$$

$$\leq \sum_{\ell=m}^N P_{n,p}(A_k | m(G) = m) P_{n,p}(m(G) = \ell) + P_{n,p}(m(G) < m)$$

$$\leq P_{n,p}(A_k | m(G) = m) \sum_{\ell=0}^N P_{n,p}(m(G) = \ell) + P_{n,p}(m(G) < m)$$

$$= P_{n,m}(A_k) + P_{n,p}(\{m(G) < m\}), \quad (16)$$

where the second inequality uses monotonicity. Similarly,

$$P_{n,p}(A_k) \geq P_{n,p}(A_k \cap \{m(G) \leq m\})$$

$$= \sum_{\ell=0}^m P_{n,p}(A_k | m(G) = \ell) P_{n,p}(m(G) = \ell)$$

$$\geq P_{n,p}(A_k | m(G) = m) \sum_{\ell=0}^m P_{n,q}(m(G) = \ell), \quad \text{using monotonicity}$$

$$= P_{n,m}(A_k) P_{n,p}(m(G) \leq m)$$

$$= P_{n,m}(A_k) - P_{n,p}(m(G) > m). \quad (17)$$
The next step is to use the following lemma, a corollary of Hoeffding’s inequality, that provides a sharp upper bound for the probability of the event that a graph $G$ in $G_{n,p}$ does not have exactly $m$ edges when $p$ is close to $m/N$. The proof of the lemma is in Appendix B.

**Lemma 5.** For $\beta$ with $0 < \beta < 1$ if $m$ is large enough and $p_1 \equiv \frac{m - m^{1+\beta}}{N}$ and $p_2 \equiv \frac{m + m^{1+\beta}}{N}$, we have

\[
\begin{align*}
\mathbb{P}_{n,p_1}(m(G) > m) &\leq e^{-m^\beta/8}, \\
\mathbb{P}_{n,p_2}(m(G) < m) &\leq e^{-m^\beta/8}.
\end{align*}
\]

Now we can use Eqs. (17) and (19) and monotonicity to obtain

\[
\begin{align*}
\mathbb{P}_{n,m}(A_k) &\leq \mathbb{P}_{n,p_2}(A_k) + \mathbb{P}_{n,p_2}(m(G) < m) \\
&\leq \mathbb{P}_{n,m}(A_k) + \mathbb{P}_{n,p_2}(m(G) < m) \\
&\leq \mathbb{P}_{n,m}(A_k) + e^{-m^\beta/N}.
\end{align*}
\]

Similarly, Eqs. (16) and (18), and monotonicity give

\[
\begin{align*}
\mathbb{P}_{n,m}(A_k) &\geq \mathbb{P}_{n,p_1}(A_k) + \mathbb{P}_{n,p_1}(m(G) > m) \\
&\geq \mathbb{P}_{n,m}(A_k) - e^{-m^\beta}.
\end{align*}
\]

Combining the above two inequalities and using Eq (15) we have,

\[
\left|1 - \frac{\mathbb{P}_{n,m}(A_k)}{\mathbb{P}_{n,m}(A_k)}\right| \leq \exp\left\{ o(1) + \sum_{r=3}^{k} |C_r| \left(\frac{m}{N}\right)^r - \frac{m^\beta}{8}\right\}. \tag{20}
\]

Next, we use $\sum_{r=3}^{k} |C_r| \left(\frac{m}{N}\right)^r = O(n^{k\alpha})$ to show that we can take $\beta$ such that $O(n^{k\alpha}) - m^\beta/8$ goes to $-\infty$. This is possible as long as $m^\beta$ grows faster than $n^{k\alpha}$. But this is possible since $\beta$ can be any constant less than 1 and $\alpha < 1/(2k - 1)$ and we are assuming that $m$ grows as fast as $n$. Combining this with Eq. (15) we have

\[
\mathbb{P}_U(G) = \frac{1}{|G_{n,m,k}|} = \frac{1}{\binom{N}{m}} \mathbb{P}_{n,m}(A_k) = \frac{1}{\binom{N}{m}} \exp\left[o(1) - \sum_{r=3}^{k} |C_r| \left(\frac{m}{N}\right)^r\right].
\]

This finishes the proof of Lemma 2.
5.3. Proof of Lemma 3

Before going into the details we will provide a high level overview of the proof of this lemma with a focus on $S_1(G)$.

A high-level preview of the proof. By definition $S_1(G) = -\sum_{t=0}^{m-1} E_{\pi} E_k(G_t^\pi, \pi(t+1))$. The first approximation we use is to change the randomness given by $\pi$. The partial graph $G_t^\pi$ is a uniformly random subgraph of $G$ that has exactly $t$ edges. Instead we look at $G_\theta$ which is a random subgraph of $G$ that has each edge of $G$ independently with probability $\theta = t/m$. The subgraph $G_\theta$ has $t$ edges in expectation which makes it a good approximation for $G_t^\pi$. We will show that via this approximation $-\sum_{t=0}^{m-1} E_{\pi} E_k(G_t^\pi, \pi(t+1))$ is approximately equal to $-m E_{\theta} \int_0^1 E_k(G_\theta, (ij)) d\theta$ where $(ij)$ is a uniformly random edge of $G$. Next, we note that $E_k$ is sum of the terms $q_r r - \ell - 1$ for all pairs $(\gamma, ij)$ where $\gamma$ is in $C_{r,\ell}$, and $(ij)$ is an edge in $(G \setminus G_\theta) \cap \gamma$. For any fixed $r, \ell$ we will show that the expected number of such $(\gamma, ij)$ pairs is dominated by the cases where $|\gamma \cap G_\theta| = |\gamma \cap G| - 1 = \ell$; in other words when $(ij)$ is the only edge of $G \cap \gamma$ that is not in $G_\theta$. Moreover $q_\ell$ is approximately equal to $(1 - \theta)m/N$. Therefore the contribution of the pairs $(\gamma, ij)$ in $S_1(G)$ for any fixed $r, \ell$ (i.e. for all $\gamma \in C_{r,\ell}$) is the same. Thus we obtain the right hand side of Lemma 3(a).

Next, we will state three axillary lemmas that will be used for the proof. But first we introduce an important subset of $\mathbb{G}_{n,m,k}$. For any graph $G$, denote its maximum degree by $\Delta(G)$. Let $\mathbb{H}_{n,m,k} \subset \mathbb{G}_{n,m,k}$ be the set of all graphs $G$ that satisfy the condition $\Delta(G) \leq n^{(k+3)\alpha}$. The next lemma will show that $\mathbb{H}_{n,m,k}$ contains almost all of $\mathbb{G}_{n,m,k}$ and its proof is given in Appendix B.

**Lemma 6.** If $m,n,k$ satisfy conditions of Lemma 3 then $|\mathbb{H}_{n,m,k}| \geq [1 - o(1)] |\mathbb{G}_{n,m,k}|$.

We also need to state the following useful upper bound, proved in Appendix B on the terms $N_{r,\ell}^{G_t^\pi, ij}$ appearing in $S_i$’s.

**Lemma 7.** If $m,n,k$ satisfy conditions of Lemma 3 then for all $3 \leq r \leq k$, $0 \leq \ell \leq r - 1$, $G \in \mathbb{H}_{n,m,k}$, and $(ij) \in G$ we have $N_{r,\ell}^{G_t^\pi, ij} = O\left(n^{r-2-\ell+(k+3)\alpha}\right)$.
Definition 2. Let \( e_1, \ldots, e_s \) be a set of \( s \) edges of \( G \). Define \( A^t_{e_1, \ldots, e_s} \) to be the event that for all \( 1 \leq i \leq s \): \( e_i \in G_t^* \). Similarly, define \( B^t_{e_1, \ldots, e_s} \) to be the event that for all \( 1 \leq i \leq s \): \( e_i \notin G_t^* \). Let also \( C^t_{e_i} \) be the event that \( \pi(t + 1) = e_i \).

Lemma 8. If \( m, n, k \) satisfy conditions of Lemma 3 then for any three positive integers \( a, b, c \) in \( [k] \) and any set of edges \( e_1, e_2, \ldots, e_{a+b+1} \) of \( G \) the following hold

\[
(a) \sum_{t=0}^{m-1} \mathbb{P} \left( A^t_{e_1, \ldots, e_a} \cap B^t_{e_{a+1}, \ldots, e_{a+b}} \right) (1 - \frac{t}{m})^c \leq O(1) + \left( m + O(1) \right) \int_0^1 \theta^a (1 - \theta)^b \, d\theta.
\]

\[
(b) \sum_{t=0}^{m-1} \mathbb{P} \left( A^t_{e_1, \ldots, e_a} \cap B^t_{e_{a+1}, \ldots, e_{a+b}} \cap C^t_{e_{a+b+1}} \right) (1 - \frac{t}{m})^c \leq O(1) + \left( 1 + O(\frac{1}{m}) \right) \int_0^1 \theta^a (1 - \theta)^b \, d\theta.
\]

\[
(c) \sum_{t=0}^{m-1} \mathbb{P} \left( A^t_{e_1, \ldots, e_a} \right) \geq -O(\frac{1}{\sqrt{m}}) + \left( m - O(\sqrt{m}) \right) \int_0^1 \theta^a \, d\theta.
\]

Proof of Lemma 8 is provided in Appendix B. Next, we prove Lemma 3.

Proof of Lemma 3 (a). Recall that

\[
S_1(G) = -\sum_{t=0}^{m-1} \sum_{r=3}^{m-1} \sum_{\ell=0}^{r-2} E_{r, \ell} N^{G_t, \pi(t+1)}_{r, \ell} \theta^{r-1-\ell}.
\]

where \( N^{G_t, \pi(t+1)}_{r, \ell} \) is number of cycles of length \( r \) in \( K_n \) that include edge \( \pi(t + 1) \) and have exactly \( \ell \) edges belonging to \( G_t \). Therefore every cycle that belongs to \( C_{r, s}(G) \) may have some contribution in \( S_1(G) \). Let us fix a cycle \( \gamma_{r, s} \in C_{r, s}(G) \). We will calculate contribution of \( \gamma_{r, s} \) in \( S_1(G) \). Let denote this contribution by \( s_1(\gamma_{r, s}) \). Let \( \{e_1, \ldots, e_s\} \) be the set of all \( s \) edges in \( \gamma_{r, s} \cap G \). In order for \( \gamma_{r, s} \) to be considered in \( N^{G_t, \pi(t+1)}_{r, \ell} \) we need to have \( \ell + 1 \) distinct indices \( i_1, \ldots, i_{\ell+1} \) in \([s]\) such that \( \{e_{i_1}, \ldots, e_{i_\ell}\} \in G_t, e_{i_{\ell+1}} = \pi(t + 1) \) and \( \{e_1, \ldots, e_s\} \setminus \{e_{i_1}, \ldots, e_{i_{\ell+1}}\} \in G \setminus (G_t \cup \{e_{\ell+1}\}) \). Therefore,

\[
s_1(\gamma_{r, s}) = -\sum_{\ell=0}^{r-2} \sum_{i_1, \ldots, i_{\ell+1}} \sum_{t=0}^{m-1} \mathbb{P}(A^t_{e_{i_1}, \ldots, e_{i_\ell}} \cap C^t_{e_{i_{\ell+1}}} \cap B^t_{e_{i_1}, \ldots, e_{i_{\ell+1}}} \setminus \{e_{i_1}, \ldots, e_{i_{\ell+1}}\}) \theta^{r-1-\ell}. \tag{21}
\]

Now, writing

\[
q_t = \frac{m - t}{N - t} = \frac{N}{N - m} \frac{m m - t}{m},
\]

and applying Lemma 8(c) and Eq. (21) yields

\[
s_1(\gamma_{r, s}) \geq - \left( \frac{N}{N - m} \right)^{r-1} \sum_{\ell=0}^{r-2} \sum_{i_1, \ldots, i_{\ell+1}} \left( \frac{m}{N} \right)^{r-1-\ell} \left[ O\left( \frac{1}{m} \right) + O\left( \frac{1}{m} \right) \int_0^1 \theta^\ell (1 - \theta)^{r+s-2\ell-2} \, d\theta \right]
\]

\[
= - e^{O(n^{a-1})} \sum_{\ell=0}^{s-1} \binom{s}{\ell} (s - \ell) \left( \frac{m}{N} \right)^{r-1-\ell} \left[ O\left( \frac{1}{m} \right) + O\left( \frac{1}{m} \right) \int_0^1 \theta^\ell (1 - \theta)^{r+s-2\ell-2} \, d\theta \right]
\]

\[
= - O\left( \frac{m^{r-s-1}}{N^{r-s}} \right) - \sum_{\ell=0}^{s-1} \binom{s}{\ell} (s - \ell) \left( \frac{m}{N} \right)^{r-1-\ell} \left[ O\left( \frac{1}{m} \right) \int_0^1 \theta^\ell (1 - \theta)^{r+s-2\ell-2} \, d\theta \right].
\]
It is easy to see that absolute value of the second term is dominated by \(1 + O(m/N)\) times its value when \(\ell = s - 1\). Hence,

\[
s_1(\gamma_{r,s}) \geq -O \left( \frac{m^{r-s-1}}{N^{r-s}} \right) - \left( \frac{m}{N} \right)^{r-s} s \left[ 1 + O \left( \frac{m}{N} \right) \right] \int_0^1 \theta^{s-1}(1-\theta)^{r-s} d\theta
\]

\[
= -O \left( \frac{m^{r-s-1}}{N^{r-s}} \right) - \left[ 1 + O(n^{\alpha-1}) \right] \left( \frac{m}{N} \right)^{r-s} s \int_0^1 \theta^{s-1}(1-\theta)^{r-s} d\theta. \tag{22}
\]

Therefore, considering all possible cycles \(\gamma_{r,s}\) we obtain

\[
S_1(G) \geq -o(1) - \left[ 1 + O \left( \frac{m}{N} \right) \right] \sum_{r=3}^{k} \sum_{s=1}^{r-1} |C_{r,s}(G)| \left( \frac{m}{N} \right)^{r-s} s \int_0^1 \theta^{s-1}(1-\theta)^{r-s} d\theta.
\]

The last step involves showing that the \(O(m/N)\) term can be removed. In particular, using \(|C_{r,s}(G)| \leq s \sum_{(ij) \in G} N_{r,s-1}^G \) and Lemma \[7\] we have

\[
O \left( \frac{m}{N} \right) \sum_{r=3}^{k} \sum_{s=1}^{r-1} |C_{r,s}(G)| \left( \frac{m}{N} \right)^{r-s} s \int_0^1 \theta^{s-1}(1-\theta)^{r-s} d\theta = O \left( \sum_{r=3}^{k} \sum_{s=1}^{r-1} mn^{r-s-1+(k+3)(s-1)+3} \left( \frac{m}{N} \right)^{r-s+1} \right)
\]

\[
= O \left( \sum_{r=3}^{k} \sum_{s=1}^{r-1} n^{\alpha[(k+3)(s-1)+(r-s)+2]-1} \right)
\]

\[
= O \left( n^{\alpha(k+3)(k-1)-1} \right)
\]

\[
= o(1).
\]

This finishes the proof of part (a).

**Proof of Lemma \[3\] (b).** First we need to approximate the number of forbidden pairs \(F_r(G^*_i)\).

\[
F_r(G^*_i) = \sum_{(ij)} \mathbb{1}(N^G_{r,r-1} > 0)
\]

\[
\geq \sum_{(ij)} N^G_{r,r-1} - \sum_{(ij)} (N^G_{r,r-1})^2
\]

\[
= |C_{r,r-1}(G^*_i)| - \sum_{(ij)} (N^G_{r,r-1})^2
\]

\[
= \sum_{\gamma_{r,r-1} \in \mathcal{C}_{r,r-1}(G^*_i)} \mathbb{1}(\gamma_{r,r-1} \in \mathcal{C}_{r,r-1}(G^*_i)) - O(n^{2k^2\alpha}),
\]

where the last step uses Lemma \[7\]. On the other hand for any cycle \(\gamma_{r,r-1} \in \mathcal{C}_{r,r-1}(G)\), using Lemma \[8\] we have

\[
\sum_{t=0}^{m-1} \mathbb{E}_z \mathbb{1}(\gamma_{r,r-1} \in \mathcal{C}_{r,r-1}(G^*_i)) \geq -O\left( \frac{1}{\sqrt{m}} \right) + (m - \sqrt{m}) \int_{\theta=0}^1 \theta^{r-1} d\theta.
\]
Thus,
\[ S_2(G) = \frac{1}{N} \sum_{r=3}^{k} \sum_{t=0}^{m-1} \mathbb{E}_\pi F_t(G_t^r) \]
\[ \geq -O\left(n^{(2k^2+1)\alpha - 1}\right) + \frac{1}{N} \sum_{r=3}^{k} |C_{r,r-1}(G)| (m - \sqrt{m}) \int_{\theta=0}^{1} \theta^{r-1} d\theta \]
\[ \geq -o(1) + O\left(\sqrt{m} \sum_{r=3}^{k} |C_{r,r-1}(G)| \int_{\theta=0}^{1} \theta^{r-1} d\theta\right) + \frac{m}{N} \sum_{r=3}^{k} |C_{r,r-1}(G)| \int_{\theta=0}^{1} \theta^{r-1} d\theta. \quad (23) \]

Similar as in the last step of proving part (a) we can use Lemma 7 to show the second term on right hand side of (23) is o(1). This concludes the proof of part (b).

Proof of Lemma 3 (c). Recall the set \(Q(G_t)\) from 23. First note that by some algebraic calculations we obtain
\[ S_3(G) = -\sum_{t=0}^{m-1} \mathbb{E}_\pi \log \left( \frac{\sum_{(ij)\in Q(G_t^r)} \exp \left( -\sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right)}{\sum_{(ij)\in Q(G_t^r)} 1} \right). \]

Now using \(e^{-x} \leq 1 - x + \frac{x^2}{2}\) for \(x > 0\) we have
\[ S_3(G) \geq -\sum_{t=0}^{m-1} \mathbb{E}_\pi \log \left( 1 - \sum_{(ij)\in Q(G_t^r)} \left( \sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right) + \frac{1}{2} \left( \sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right)^2 \right). \]

Now using \(-\log(1 - x) \geq x\) for \(x > 0\) and \(|Q(G_t^r)| \leq N\) the last term can be simplified to obtain
\[ S_3(G) \geq \mathbb{E}_\pi \left[ \frac{1}{N} \sum_{t=0}^{m-1} \sum_{(ij)\in Q(G_t^r)} \left( \sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right) - \frac{1}{2N} \sum_{t=0}^{m-1} \sum_{(ij)\in Q(G_t^r)} \left( \sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right)^2 \right]. \]

Next step is similar to the last part of proof of part (a) which uses Lemma 7 to give us
\[ S_3(G) \geq \frac{1}{N} \sum_{t=0}^{m-1} \mathbb{E}_\pi \left[ \sum_{(ij)\in Q(G_t^r)} \left( \sum_{r=3}^{k} \sum_{t=0}^{m-2} N_{r,t}^{G_t^r,ij} q_t^{r-1}\right) \right] - O(n^{2k(k+3)\alpha - 1}). \quad (24) \]

Now, similar to the proof of (a) we will find contribution of a cycle \(\gamma_{r,s} \in C_{r,s}(G)\) that is denoted by
\[ s_3(\gamma_{r,s}). \] The only difference is that this time the edge \((ij) = \pi(t + 1)\) could be also outside of \(G\). In other words, it can be any of the \(r - s\) edges in \(\gamma_{r,s} \setminus \{e_1, \ldots, e_s\}\). Depending on whether \((ij)\) belongs to \(G\) or not, the summation breaks into two pieces. Then we use parts (a) and (b) of Lemma 8 respectively. Similar to Eq. (22), we obtain
\[ s_3(\gamma_{r,s}) \geq -O \left( \frac{m^{r-s-1}}{N^{r-s}} \right) - \left( \frac{m}{N} \right)^{r-s} \int_{0}^{1} \theta^{s} (1 - \theta)^{r-s-1} d\theta. \]
Therefore,
\[
S_3(G) \geq -o(1) - \sum_{r=3}^{k} \sum_{s=0}^{r-1} |C_{r,s}(G)| \left( \frac{m}{N} \right)^{r-s} (r-s) \int_0^1 \theta^r (1-\theta)^{r-s-1} d\theta
\]
which finishes the proof □

6. Running Time of RandGraph and Proof of Theorem 2

In this section we will prove that RandGraph can be implemented in a way that its expected running time would be \(O(n^2 m)\) operations. The idea is to define a surrogate quantity for probabilities \(p(ij|G_t)\) that are efficiently computable using sparse matrix multiplication of order \(O(n^2)\) per each step of the algorithm. The key point is that in the definition of \(p(ij|G_t)\) in [3] we the number of certain simple cycles. It is known that one can count number of (not necessarily) simple cycles of a graph via matrix multiplication of the its adjacency matrix. We will use this fact and prove that the contribution of non-simple cycles will be negligible.

This method counts non-simple cycles as well. This was done to simplify the implementation of the algorithm. For the analysis it is convenient to work with simple cycles and the above lemma provides the necessary comparison between these two methods of counting cycles.

During the execution of RandGraph, after adding \(t\) edges, let \(M_t\) and \(M_t^{(c)}\) be the adjacency matrices of the partially constructed graph \(G_t\) and its complement \(G_t^{(c)}\) respectively. In addition, let \(Q_t\) be the adjacency matrix of the graph obtained by all edges \((ij)\) such that \(G_t \cup (ij) \in \mathcal{G}_{n,t+1,k}\).

We modify RandGraph so that it selects the \((t+1)\)th edge from all pairs \((ij)\) such that \(G_t \cup (ij) \in \mathcal{G}_{n,t+1,k}\). Let us call this modification RandGraph’. The key result of this section is the following Lemma and is proved in Appendix B.
Lemma 9. For any non-zero probability term \( p'(ij|G_t) \),

\[
p'(ij|G_t) \geq \frac{1}{Z(G_t)} e^{-E_k(G_t,ij)} - O\left(n^{k(k+3)\alpha/2}\right),
\]

where \( Z(G_t) = \sum_{rs \in Q(G_t)} e^{-E_k(G_t,rs)} \) is the normalization term in definition of \( p(ij|G_t) \) from §3.

Using Lemma 9, we can see that the output distribution of \( \text{RandGraph}' \) still satisfies the inequality \( P_{RG'}(G) \geq (1 - o(1))P_U(G) \) for all but \( o(|G_{n,m,k}|) \) graphs \( G \) in \( G_{n,m,k} \). More formally, a variant of Lemma 1 holds for \( P_{RG'} \) using Lemma 1 for \( P_{RG} \) and Lemma 9. Next, we focus on the implementation of \( \text{RandGraph}' \).

The fact that \( \text{RandGraph}' \) has polynomial running time is clear since the matrix of the probabilities at any step, \( P_{G_t} \), can be calculated using matrix multiplication. In fact a naïve calculation shows that \( P_{G_t} \) can be calculated with \( O(kn^3) \) operations. This is because \( r^{th} \) power of a matrix for any \( r \) takes \( O(rn^3) \) operations to compute. So we obtain the simple bound of \( O(n^3m) \) for the running time. But we can improve this running time by at least a factor \( n \) with exploiting the structure of the matrices.

Notice that the adjacency matrix \( Q_t \) is equal to \( J_n - \widehat{\text{sign}}(\sum_{r=0}^{k-1} M_r^t) \) where \( J_n \) is the \( n \) by \( n \) matrix of all ones and the \( \widehat{\text{sign}}(B) \) for any matrix \( B \) means the “sign” function is applied to each entry of \( B. \) This is correct since any “bad” pair \( (ij) \), that cannot be added to \( G_t \), corresponds to a path in \( G_t \) of length \( r \) between \( i \) and \( j \) for \( 0 \leq r \leq k - 1 \). Such path forces the \( ij \) entry of the matrix \( M_r^t \) to be positive.

Now we can store the matrices \( M_1, \ldots, M_t^{k-1} \) and \( M_2, \ldots, M_t^{k-1} \) at the end of each iteration and use them to efficiently calculate \( M_{t+1}, \ldots, M_{t+1}^{k-1} \) and \( M_{t+1}^2, \ldots, M_{t+1}^{k-1} \). This is because the differences \( M_{t+1} - M_t \) and

\[
\left[ M_{t+1} + \frac{m - t + 1}{\binom{n}{2} - t + 1} M_{t+1}^{(c)} \right] - \left[ M_t + \frac{m - t}{\binom{n}{2} - t} M_t^{(c)} \right]
\]

are sparse matrices and updating the matrix multiplications can be done with \( O(n^2) \) operations which reduces the overall running time to \( O(n^2m) \). More precisely, we can use

\[
M_{t+1}^r = [M_t + (M_{t+1} - M_t)]^r = M_t^r + L,
\]
where \( L \) is a linear sum of matrix products where each term contains at least one of \((M_{t+1} - M_t), \ldots, (M_{t+1} - M_t)^{r-1}\). Since \( M_{t+1} - M_t \) has \( O(1) \) non-zero entries then the total operations required for calculating \( L \) is of \( O(n^2) \). A similar argument can be used for calculating

\[
\left[ M_{t+1} + \frac{m-t+1}{(n/2)-t+1} M_{c(t+1)} \right]^r
\]

using sparsity of both \( M_{t+1} - M_t \) and \( M_{c(t+1)} - M_{c(t)} \).

Since Theorem 1 shows that \( \text{RandGraph} \) and hence \( \text{RandGraph}' \) are successful with probability \( 1 - o(1) \), the expected running-time of \( \text{RandGraph}' \) for generating an element of \( G_{n,m,k} \) is \( O(n^2m) \) as well which finishes proof of Theorem 2.

7. Extension to Bipartite Graphs with Given Degrees

The ideas described in §4 can be used to generate random bipartite graphs with given node degrees. Such graphs define the standard model for irregular LDPC codes. In this section we will show how to modify \( \text{RandGraph} \) for this application. The analysis of this extension is somewhat cumbersome and is beyond the scope of this paper but we expect it to be conceptually similar to the analysis of \( \text{RandGraph} \). Since this is a short section, the notation introduced here is not presented in Table 1.

Consider two sequences of positive integers \( \bar{r} = r_1, \ldots, r_{n_1} \) and \( \bar{c} = c_1, \ldots, c_{n_2} \) for degrees of the vertices such that \( m = \sum_{i=1}^{n_1} r_i = \sum_{j=1}^{n_2} c_j \). We would like to generate a random bipartite graph \( G(V_1, V_2) \), \( V_1 = [n_1] \) and \( V_2 = [n_2] \), with girth greater than \( k \) and with degree sequence \( (\bar{r}, \bar{c}) \). We also assume that \( k \) is an even number. Denote the set of all such graphs by \( G_{\bar{r}, \bar{c}, k} \). The algorithm is a natural generalization of \( \text{RandGraph} \) where the probabilities \( p(ij|G_t) \) are adjusted properly. Here each probability \( p''(ij|G_t) \) is an approximation to the probability that a uniformly random extension of graph \( G_t \cup (ij) \) has girth larger than \( k \) (the intuitive reason for this is described in §4).

The estimation procedure for \( p''(ij|G_t) \) is slightly more involved than the one used for \( p(ij|G_t) \). It relies on considering a \emph{configuration model} representation for the graphs with degree sequence \( (\bar{r}, \bar{c}) \) and using a similar argument to the one in §4 to obtain the following Poisson approximation for \( p''(ij|G_t) \),

\[
p''(ij|G_t) \approx \frac{\hat{r}_i \hat{c}_j e^{-E''_{k}(G_t,ij)}}{Z''(G_t)},
\]
Algorithm 2 BipRandGraph.

Input: Degree sequence $(\bar{r}, \bar{c})$ and $k$

Output: An element of $G_{\bar{r},\bar{c},k}$ or FAIL

set $G_0$ to be a graph over vertex sets $V_1 = [n_1]$, $V_2 = [n_2]$ and with no edges.

let $\hat{r} = \{\hat{r}_1, \ldots, \hat{r}_n\}$ and $\hat{c} = \{\hat{c}_1, \ldots, \hat{c}_m\}$ be ordered sets that are initialized by $\hat{r} = \bar{r}$ and $\hat{c} = \bar{c}$

for each $t$ in $\{0, \ldots, m - 1\}$ do

if adding any edge to $G_t$ creates a cycle of length at most $k$ then

stop and return FAIL

else

sample an edge $(ij)$ from $V_1 \times V_2$ with probability $p''(ij|G_t)$, defined by Eq. (26) below

set $G_{t+1} = G_t \cup (ij)$

set $\hat{r}_i = \hat{r}_i - 1$ and $\hat{c}_j = \hat{c}_j - 1$

end if

end for

if the algorithm does not FAIL before $t = m - 1$ then

return $G_m$

end if


where $Z''(G_t)$ is a normalization term, and $\hat{r}_i$, $\hat{c}_j$, denote the remaining degrees of $i$ and $j$. Furthermore,

$$E_k''(G_t, ij) \equiv \sum_{r=1}^{k/2} \sum_{\gamma \in C_{2r}} p^t_{ij}(\gamma),$$

where $C_{2r}$ is the set of all simple cycles of length $2r$ in the complete bipartite graph on vertices of $V_1$ and $V_2$, and $p^t_{ij}(\gamma)$ is approximately the probability that $\gamma$ is in a random extension of $G_t$ to a random bipartite graph with degree sequence $(\bar{r}, \bar{c})$. More precisely,

$$p^t_{ij}(\gamma) = \frac{(m - t - 2r + |\gamma \cap G_t|)!(\prod_{\ell \in \gamma \cap V_1} R^t_{ij}(\ell, \gamma) \prod_{\ell \in \gamma \cap V_2} C^t_{ij}(\ell, \gamma)}{(m - t - 1)!}.$$
where
\[
R_{ij}^t(\ell, \gamma) = \begin{cases} 
\hat{r}_\ell(\hat{r}_\ell - 1) & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 0, \\
\hat{r}_\ell & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 1, \\
1 & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 2.
\end{cases}
\]

Similarly,
\[
C_{ij}^t(\ell, \gamma) = \begin{cases} 
\hat{c}_\ell(\hat{c}_\ell - 1) & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 0, \\
\hat{c}_\ell & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 1, \\
1 & \text{if } \deg_\ell(\gamma \cap [G_t \cup (ij)]) = 2.
\end{cases}
\]

Here the notation \(\deg_v(H)\) for a node \(v\) of graph \(G\) and subgraph \(H\) of \(G\) refers to the induced degree of \(v\) in \(H\).

**Remark 2.** The evaluation of \(p''(ij|G_t)\) appears rather complex from the above description but a fast implementation is described in [Bayati et al. 2009].

**Remark 3.** Besides cycles, other similar substructures have been suggested to be responsible for the decoding errors at high signal-to-noise ratio. Our algorithm can be adapted to exclude these substructures if one can algebraically simplify \(p''(ij|G_t)\) so that it would be implementable.

**Acknowledgments**

The authors would like to thank Balaji Prabhakar, Joel Spencer, Daniel Spielman, and anonymous referees for useful discussions and suggestions.

**References**


**Appendix A: Mathematical Notations**

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
</table>
| $[n]$    | When $n$ is a positive integer it denotes the set $\{1,2,\ldots,n\}$.
| $K_n$    | Complete graph with vertex set $[n]$.
| $O$      | For sequences $\{a_n\}_{n \geq 1}, \{b_n\}_{n \geq 1}$ big $O$ notation $a_n = O(b_n)$ means $\lim\sup_{n \to \infty} a_n/b_n < \infty$.
| $o$      | For sequences $\{a_n\}_{n \geq 1}, \{b_n\}_{n \geq 1}$ little $O$ notation $a_n = o(b_n)$ means $\lim\sup_{n \to \infty} a_n/b_n = 0$.
| $(ij)$   | An edge that connects node $i$ to node $j$ ($i,j \in [n]$) (in a graph $G$ with vertices $[n]$).
| $n$      | Number of vertices of graphs considered in the paper.
| $m$      | Number of edges of most graphs in the paper.
| $N$      | Defined to be $\binom{n}{2}$.
$m(G)$: Number edges of a graph $G$.

$\mathbb{G}_{n,m}$: Set of all simple graphs with $m$ edges and vertices $[n]$.

$\mathbb{G}_{n,p}$: Random graph model of simple graphs on $[n]$ where each edge is present (independently) with probability $p$.

$\mathbb{P}_{n,m}$: Uniform probability distribution over $\mathbb{G}_{n,m}$.

$\mathbb{P}_{n,p}$: Probability distribution obtained by random graph model $\mathbb{G}_{n,p}$.

$\mathbb{G}_{n,m,k}$: The subset of graphs in $\mathbb{G}_{n,m}$ with girth greater than $k$.

$\mathbb{H}_{n,m,k}$: The set of graphs $G$ in $\mathbb{G}_{n,m,k}$ with maximum degree of order $O(n^{k+3} \alpha)$.

$\mathbb{G}_{n,m,k}(\tau)$: Subset of graphs $G$ in $\mathbb{G}_{n,m,k}$ where $\mathbb{P}_{\mathbb{RG}}(G) < (1 - \tau) \mathbb{P}_U(G)$.

$\mathbb{P}_{\mathbb{RG}}$: Output distribution of $\text{RandGraph}$ which is a distribution on $\mathbb{G}_{n,m,k}$.

$\mathbb{P}_U$: Uniform distribution on $\mathbb{G}_{n,m,k}$.

$G_t$: Partially constructed graph in $\text{RandGraph}$ after $t$ steps.

$q_t$: Equals to $(m - t)/(N - t)$.

$\theta$: Equals to $t/m$.

$\pi$: A permutation of the edges of $G$ where $G \in \mathbb{G}_{n,m}$.

$G_t^\pi$: The graph having $[n]$ as vertex set and $\{\pi(1), \ldots, \pi(t)\}$ as edge set.

$E_\pi$: Expectation with respect to a uniformly random permutation $\pi$.

$P_\pi$: Probability with respect to a uniformly random permutation $\pi$.

$\gamma$: Notation used for cycles.

$Q(G_t)$: The set of edges $(ij)$ that do not belong to $G_t$ and $G_t \cup (ij) \in \mathbb{G}_{n,t+1,k}$.

$p(ij|G_t)$: For each $(ij) \in Q(G_t)$, it is the probability of selecting $(ij)$ in step $t$ of $\text{RandGraph}$.

$E_k(G_t, ij)$: Equals to $\sum_{r=3}^k \sum_{\ell=0}^{r-2} N_{r,\ell}^{G_t,ij} q_{r-1-\ell}$.

$T$: Execution tree of a sequential graph generation algorithm like $\text{RandGraph}$ (see §4 for details).

$\pi_t$: For a partially constructed graph $G_t$, it is an ordering (permutation) of its edges.

$n_k(G_t, \pi_t)$: Number of cycles of length at most $k$ in a random extension of of a pair $(G_t, \pi_t)$ in $T$.

$N_{r,\ell}^{G_t,ij}$: Number of simple cycles in $K_n$ that have length $r$, include $(ij)$, and include exactly $\ell$ edges of $G_t$.

$Z(G)$: Normalization constant in definition of $p(ij|G_t)$ in Eq. (1).

$Z_0(G)$: Is equal to $N - t - \sum_{r=3}^k F_r(G_t^\pi)$.

$F_r^{G_t^\pi}$: Number of edges $(ij)$ where $G_t^\pi \cup (ij)$ has a cycle of length $r$.

$S_1(G)$: Equals to $-\sum_{t=0}^{m-1} E_\pi E_k(G_t^\pi, \pi(t+1))$.

$S_2(G)$: Equals to $\frac{1}{N} \sum_{r=3}^k \sum_{t=0}^{m-1} E_\pi F_r(G_t^\pi)$. 
\[ S_3(G) : \text{Equals to } - \sum_{t=0}^{m-1} \mathbb{E} \log \frac{Z(G_t^+)}{Z_{st}}. \]

\( \mathcal{C}_r : \) Set of all simple cycles of length \( r \) in \( K_n \).

\( \mathcal{C}_{r,\ell}(G) : \) Cycles in \( \mathcal{C}_r \) that include exactly \( \ell \) edges of \( G \).

\( \gamma_{r,s} : \) An element of \( \mathcal{C}_{r,\ell}(G) \).

\( s_i(C_{r,s}) : \) For each \( i = 1, 2, 3 \) denotes contribution of cycle \( C_{r,s} \) in \( S_i(G) \).

\( A_k : \) The event that a random graph has girth greater than \( k \).

\( \text{deg}_v(H) : \) Induced degree of a note \( v \) in a subgraph \( H \) of a larger graph containing \( v \).

\( \Delta(G) : \) Maximum degree of graph \( G \).

\( A_{e_1,\ldots,e_s}^{\pi,\ell} : \) The event \( \{ \forall i \in [s] : e_i \in G_t^\pi \} \) when \( e_1,\ldots,e_s \) are edges of \( G \).

\( B_{e_1,\ldots,e_s}^{\pi} : \) The event \( \{ \forall i \in [s] : e_i \notin G_t^\pi \} \) when \( e_1,\ldots,e_s \) are edges of \( G \).

\( C_e^{\pi} : \) The event \( \{ \pi(t+1) = e \} \) for edge \( e \) in \( G \).

\( M_t : \) Adjacency matrix of \( G_t \).

\( M_t^{\pi} : \) Adjacency matrix of complement of \( G_t \).

\( Q_t : \) Adjacency matrix of all edges in \( Q(G_t) \).

\( \mathbf{A} = \mathbf{B} \odot \mathbf{C} : \) For \( n \times n \) matrices \( \mathbf{A, B, C} \) it means that for all \( i,j \in [n] : a_{ij} = b_{ij}c_{ij} \).

\( \mathbf{A} = \exp(\mathbf{B}) : \) For \( n \times n \) matrices \( \mathbf{A, B} \) it means that for all \( i,j \in [n] : a_{ij} = e^{b_{ij}} \).

\( \mathbf{A} = \text{sign}(\mathbf{B}) : \) For \( n \times n \) matrices \( \mathbf{A, B} \) it means that for all \( i,j \in [n] : a_{ij} = \text{sign}(b_{ij}) \).

\( J_n : \) It is the \( n \) by \( n \) matrix of all ones.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
\textbf{Notation} & \textbf{Description} \\
\hline
\text{deg}_v(H) & Induced degree of a note \( v \) in a subgraph \( H \) of a larger graph containing \( v \). \\
\hline
\Delta(G) & Maximum degree of graph \( G \). \\
\hline
\end{tabular}
\caption{Mathematical notations.}
\end{table}

\section*{Appendix B: Proofs of Axillary Lemmas}

\textbf{Proof of Lemma 4} \hspace{1em} It is easy to see that \( |\mathcal{C}_r| = \text{constant} \cdot n^r \). Now we try to find an upper bound for the number of paths in \( \mathcal{C}_r \setminus \mathcal{C}_{r,0}(G) \) which is the set of paths of length \( r \) that intersect at least one edge of \( G \). The number of paths \( \gamma \) that intersect a fixed edge \( (ij) \) in \( G \) is of order \( O(n^{r-2}) \) since there are \( \binom{n-2}{r-2} \) ways to pick the remaining \( r-2 \) vertices of \( \gamma \) and this is the dominating term. And Therefore,

\[ \frac{|\mathcal{C}_r \setminus \mathcal{C}_{r,0}(G)|}{|\mathcal{C}_r|} = O \left( \frac{\sum_{i,j \in G} n^{r-2}}{n^r} \right) = O \left( mn^{-2} \right) = O \left( n^{\alpha-1} \right) \quad \square \]

\textbf{Proof of Lemma 5} \hspace{1em} First we state the following modified version of Hoeffding’s inequality, adapted from Corollary 3.2 in Steger and Wormald (1999).
**Proposition 2 (Hoeffding's inequality).** Let $X_1, \ldots, X_n$ be independent variables with $0 \leq X_i \leq 1$ for all $i \in [n]$, and let $X = \sum_{i=1}^{n} X_i$. Then for $\delta \leq 4/5$,

$$
P \left[ \left| X - \mathbb{E}(X) \right| > \delta \mathbb{E}(X) \right] \leq e^{-\delta^2 \mathbb{E}(X)/4}.
$$

We can now take $N$ iid Bernoulli($p$) random variables corresponding to the edges of $G$ in $\mathbb{G}_{n,p}$ and use Proposition 2 to obtain, for any $0 < p < 1$ and $0 < \delta < 4/5$,

$$
P_{n,p} \left( \left| m(G) - Np \right| > \delta Np \right) \leq e^{-\delta^2 Np/4}.
$$

Now we can see that by taking $\delta = \frac{m(1+\beta/2)}{m+m(1+\beta/2)}$, when $\beta \in (0,1)$ and $m$ is large enough, we have $\delta < 4/5$, $(1 + \delta)Np_1 = m$, and $\delta^2 Np_1 \geq m^\beta/2$ which give

$$
P_{n,p_1} \left( m(G) > m \right) \leq P_{n,p_1} \left( m(G) > (1 + \delta)Np_1 \right) 
\leq e^{-\delta^2 Np_1/4} 
\leq e^{-m^\beta/8}.
$$

For the second inequality, $P_{n,p_2} \left( m(G) < m \right) \leq e^{-m^\beta/8}$, we take $\delta = \frac{m(1+\beta/2)}{m+m(1+\beta/2)}$, which gives $(1 - \delta)Np_2 = m$ and $\delta^2 Np_2 \geq m^\beta/2$ for large $m$ and the result similarly follows $\square$

**Proof of Lemma 6** We use the following version of Chernoff inequality, Theorem A.1.18 of page 170 in (Alon and Spencer 1992). For i.i.d. Bernoulli random variables $X_1, \ldots, X_N$ with mean $p$

$$
P \left( \sum_{i=1}^{N} X_i > \eta + Np \right) < e^{-2\eta^2}.
$$

Now combining the above inequality with union bound, for graphs $G$ in $\mathbb{G}_{n,m,k}$ we have for any $p \in (0,1)$

$$
P_{n,p} \left[ \Delta(G) > (n-1)p + \eta \right] < ne^{-2\eta^2}.
$$

Note that the event $\{ \Delta(G) > (n-1)p + \eta \}$ is a monotone property (see beginning of §5.2.1 for definition). Therefore, similar to the proof of Lemma 2 we can take $p_2 = \frac{m+m(1+\beta/2)}{N}$ and use (17) and (19) to obtain

$$
P_{n,m} \left[ \Delta(G) > (n-1)p_2 + \eta \right] < P_{n,p_2} \left[ \Delta(G) > (n-1)p_2 + \eta \right] + P_{n,p_2} \left[ m(G) < m \right] 
< ne^{-2\eta^2} + e^{-m^\beta/8}.
$$

Thus, for $\beta = 0.5$ and $\eta = n^{(k+2)\alpha/2}$, combining the above bounds with $np_2 = O(n^\alpha)$ and $m^\beta/8 > 2n^{(k+2)\alpha}$ we have

$$
P_{n,m} \left[ \Delta(G) > n^{(k+3)\alpha} \right] < e^{-n^{(k+1)\alpha}}.
$$

(27)
Finally, note that in [§5.2] we explicitly calculated $\mathbb{P}_{n,m}(A_k)$ which shows that $\mathbb{P}_{n,m}(A_k)^{-1}$ is of order $e^{O(n^{k\alpha})}$. Hence,

$$\frac{\mathbb{H}_{n,m,k}}{\mathbb{G}_{n,m,k}} = \mathbb{P}_{n,m}(\Delta(G) \leq n^{(k+3)\alpha} | G \in \mathbb{G}_{n,m,k})$$

$$= \mathbb{P}_{n,m}(\Delta(G) \leq n^{(k+3)\alpha} \cap A_k)$$

$$= \mathbb{P}_{n,m}(A_k) - \mathbb{P}_{n,m}(\Delta(G) > n^{(k+3)\alpha})$$

$$> 1 - e^{-n^{(k+1)\alpha} + O(n^{\kappa\alpha})} > 1 - o(1).$$

This finishes proof of Lemma 6 □

Proof of Lemma 7 Clearly $N_{r,t}^{G_t,i,j}$ is bounded from above by the number of paths (not necessarily simple paths) of length $r-1$ from $i$ to $j$ that have at least $\ell$ edges of the $G_t$. Number of all such paths is equal to the number of sequences $C = (i = i_0, i_1, \ldots, i_{r-1} = j)$ with $i_s \in [n]$ for all $s$, and at least $\ell$ of consecutive pairs $(i_s, i_{s+1})$ in $G_t$. Since $\ell < r - 1$ there is a pair $(i_s, i_{s+1})$ that does not belong to $G_t$. We take $s$ to be the smallest such number. So any path $C$ breaks into $C = C_1 \cup C_2$ where $C_1$ is a path starting from $i$ with length $s$ and completely lies inside $G_t$. Number of such paths is at most $\Delta(G)^s$. Similarly $C_2$ is a path with one endpoint equal to $j$ and length $r - 2 - s$ that contains $\ell - s$ edges of $G_t$. Number of such paths is at most $\Delta(G)^{\ell - s} n^{r-2-\ell}$. Therefore using $G \in \mathbb{H}_{n,m,k}$,

$$N_{r,t}^{G_t,i,j} \leq \sum_{s=0}^{t} \Delta(G)^f n^{r-2-\ell} = O(n^{r-2-\ell+(k+3)\alpha}),$$

(28)

which finishes the proof □

Proof of Lemma 8 Note that $G_t^\pi$ is a random subgraph of $G$ that has $t$ edges. Therefore

$$\mathbb{P}_{\pi}(A_{i_1 \ldots i_a} \cap B_{e_{a+1} \ldots e_{a+b}}^\pi) = \left(\frac{m-a-b}{m}\right) \left(\frac{m-a}{m}\right)^A \left(\frac{m-a-b}{m}\right)^{A-s} \left(\frac{m}{m}\right)^{B-s}$$

$$= \left[\frac{m^{a+b}}{m^{a-b+1}}\right] \left(\frac{m-t}{m-t-b+1}\right)^{A} \left(\frac{m-t-b}{m-t}\right)^{B}$$

$$= \left[\frac{t^{a-b+1}}{t^a}\right] f_{a,b}(t)$$

where $f_{a,b}(t) = \left(\frac{t}{m}\right)^a \left(\frac{m-t}{m}\right)^b$. Therefore:

$$\mathbb{P}_{\pi}(A_{i_1 \ldots i_a} \cap B_{e_{a+1} \ldots e_{a+b}}^\pi) (1 - \frac{t}{m})^c \leq (1 + \frac{a+b}{m-a-b}) f_{a,b+c}(t)$$

$$\leq \left(1 + O\left(\frac{1}{m}\right)\right) f_{a,b+c}(t).$$

(29)

Now using the fact that the function $\theta^a (1 - \theta)^b$ has at most one maximum in the interval $(0, 1)$ then

$$\sum_{t=0}^{m-1} f_{a,b+c}(t) \leq \int_{\theta=0}^{1} \theta^a (1 - \theta)^{b+c} d\theta + O\left(\frac{1}{m}\right).$$

(30)
Combining Eqs. (29) and (30) proves part (a) of Lemma 8.

Part (b) is now easy to prove to the above bound and the following

\[ \Pr \left( A_{e_1,\ldots,e_a}^t \cap B_{e_{a+1},\ldots,e_{a+b}}^t \cap C_{e_{a+b+1}}^t \right) (1 - \frac{t}{m})^c = \frac{(m-a-b)}{(m-a-b)} \left( 1 - \frac{t}{m} \right)^c \leq O \left( \frac{1}{m} \right) \left( 1 + O \left( \frac{1}{m} \right) \right) f_{a,b+c}(t). \]

Hence, we are left with proof of part (c). First we use Bernoulli’s inequality \((1-x)^y \geq 1 - yx\) for \(0 < x, y \geq 1\) to show that for \(t > \sqrt{m}\)

\[ \Pr \left( A_{e_1,\ldots,e_a}^t \right) = \frac{(m-a)}{(m)} \geq (1 - \frac{a}{t})^a f_{a,0}(t) \geq \left( 1 - O \left( \frac{1}{\sqrt{m}} \right) \right) f_{a,0}(t). \quad (31) \]

Then, using that \(f_{a,0}(t)\) is increasing in \(t\) we obtain

\[ \frac{\sum_{t=0}^{m-1} f_{a,0}(t)}{m} \geq \int_{\theta=0}^{1} \theta^a d\theta - O \left( \frac{1}{m} \right). \quad (32) \]

Hence,

\[ \sum_{t=0}^{m-1} \Pr \left( A_{e_1,\ldots,e_a}^t \right) \geq \sum_{t > \sqrt{m}} \Pr \left( A_{e_1,\ldots,e_a}^t \right) \geq \left( 1 - O \left( \frac{1}{\sqrt{m}} \right) \right) \sum_{t > \sqrt{m}} f_{a,0}(t) \geq \left( 1 - O \left( \frac{1}{\sqrt{m}} \right) \right) \sum_{t=0}^{m-1} f_{a,0}(t) - O \left( \frac{1}{m} \right) \geq \left( m - O(\sqrt{m}) \right) \int_{\theta=0}^{1} \theta^a d\theta - O \left( \frac{1}{m} \right), \]

which finishes proof of Lemma 8. \(\square\)

**Proof of Lemma 9.** The main idea is that each entry of the matrix \(M_t + \frac{m-t}{2} \mathbf{M}_t^{(c)}\) corresponds to a sum of all products of entries of the matrix \(M_t + \frac{m-t}{2} \mathbf{M}_t^{(c)}\) that correspond to paths of length \(r\) in \(K_n\). Moreover the sum is dominated by those products that correspond to simple paths rather than self intersecting paths. Below, we will show this formally.

By definition, for any non-zero \((ij)\) entry of the matrix \(P'_{G_t}\) we have:

\[ (P'_{G_t})_{ij} = \exp \left( - \sum_{r=2}^{k-1} \sum_{\ell=0}^{r-1} N_{r+1,\ell}^{G_t} q_t^{r-\ell} - \sum_{r=2}^{k-1} \sum_{\ell=0}^{r-1} M_{r+1,\ell}^{G_t} q_t^{r-\ell} \right) \]

\[ = \exp \left( - \sum_{r=3}^{k-2} \sum_{\ell=0}^{r-1} N_{r,\ell}^{G_t} q_t^{r-1-\ell} - \sum_{r=3}^{k-2} \sum_{\ell=0}^{r-1} M_{r,\ell}^{G_t} q_t^{r-1-\ell} \right) \]

where \(M_{r,\ell}^{G_t,ij}\) is the number of self intersecting cycles of length \(r\) in \(K_n\) that include \((ij)\) and exactly \(\ell\) edges of \(G_t\). Similarly to the argument used in Lemma 7 to prove an upper bound for \(N_{r,\ell}^{G_t,ij}\), we can show that

\[ M_{r,\ell}^{G_t,ij} = O(n^{r-3-\ell+2(\ell+3)\theta}). \quad (33) \]
Therefore,

$$\left( \mathbf{P}'_{G_t} \right)_{ij} = \exp \left( - \sum_{r=3}^{k} \sum_{\ell=0}^{r-2} N_{G_t, \alpha+1, \ell}^{G_t, (ij)} q_{t}^{r-1-\ell} - O \left( n^{k(k+3)\alpha-2} \right) \right).$$

For simplicity of the notation let: $D_{ij}^{G_t} = \exp \left( - \sum_{r=3}^{k} \sum_{\ell=0}^{r-2} N_{G_t, \alpha+1, \ell}^{G_t, (ij)} q_{t}^{r-1-\ell} \right)$. Hence,

$$p'(ij|G_t) = \frac{(\mathbf{P}'_{G_t})_{ij}}{Z'(G_t)} = \frac{(\mathbf{P}'_{G_t})_{ij}}{\sum_{rs \in Q(G_t)} (\mathbf{P}'_{G_t})_{rs}}$$

$$= \frac{D_{ij}^{G_t} \exp \left( - O \left( n^{k(k+3)\alpha-2} \right) \right)}{\sum_{rs \in Q(G_t)} D_{rs}^{G_t} \exp \left( - O \left( n^{k(k+3)\alpha-2} \right) \right)} \geq \frac{D_{ij}^{G_t} \exp \left( - O \left( n^{k(k+3)\alpha-2} \right) \right)}{\sum_{rs \in Q(G_t)} D_{rs}^{G_t} \exp \left( - O \left( n^{k(k+3)\alpha-2} \right) \right)}$$

which finishes the proof $\square$